101 Rec'd PCT/PTO 14 MAY 1998

SUBSTITUTE FORM PTO-1390 U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE **TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371**		ATTORNEY'S DOCKET NUMBER 00786/292002 U.S. APPLICATION NO (IF KNOWN) 09/068804
INTERNATIONAL APPLICATION NO. PCT/US96/18504	INTERNATIONAL FILING DATE 14 November 1996	PRIORITY DATE CLAIMED 14 November 1995

TITLE OF INVENTION
SALMONELLA SECRETED PROTEINS AND USES THEROF

APPLICANT(S) FOR DO/EO/US
SAMUEL I. MILLER

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

1. ■ This is a FIRST submission of items concerning a filing under 35 U.S.C. 371.
2. □ This is a SECOND or SUBSEQUENT submission of items concerning a filing under 35 U.S.C. 371.
3. ■ This is an express request to begin national examination procedures (35 U.S.C. 371(f)) at any time rather than delay examination until the expiration of the applicable time limit set in 35 U.S.C. 371(b) and PCT Articles 22 and 39(1).
4. ■ A proper Demand for International Preliminary Examination was made by the 19th month from the earliest claimed priority date.
5. ■ A copy of the International Application as filed (35 U.S.C. 371(c)(2))
   a. □ is transmitted herewith (required only if not transmitted by the International Bureau).
   b. ■ has been transmitted by the International Bureau (and a copy is enclosed).
   c. □ is not required, as the application was filed in the United States Receiving Office (RO/US).
6. □ A translation of the International Application into English (35 U.S.C. 371(c)(2)).
7. ■ Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3))
   a. □ are transmitted herewith (required only if not transmitted by the International Bureau).
   b. □ have been transmitted by the International Bureau.
   c. □ have not been made; however, the time limit for making such amendments has NOT expired.
   d. ■ have not been made and will not be made.
8. □ A translation of amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).
9. ■ An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)) (unsigned).
10. □ A translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)).

Items 11. to 16. below concern other documents or information included:

11. □ An Information Disclosure Statement under 37 CFR 1.97 and 1.98.
12. □ An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.
13. ■ A FIRST preliminary amendment.
    □ A SECOND or SUBSEQUENT preliminary amendment.
14. □ A substitute specification.
15. □ A change of power of attorney and/or address letter.
16. ■ Other items or information: Formal Drawings

"EXPRESS MAIL" Mailing Label Number EM529138070US

Date of Deposit 14 May 1998

I hereby certify under 37 CFR 1.10 that this correspondence is being deposited with the United States Postal Service as **"Express Mail Post Office To Addressee"** with sufficient postage on the date indicated above and is addressed to the Box PCT, Assistant Commissioner for Patents, Washington, D.C. 20231.

Lisa G Gray

Lisa G Gray

U.S. APPLICATION NO. (IF KNOWN)	INTERNATIONAL APPLICATION NO.	ATTORNEY'S DOCKET NUMBER
	PCT/US96/18504	00786/292002

17. ■ The following fees are submitted:	CALCULATIONS	PTO USE ONLY
Basic National Fee (37 CFR 1.492(a)(1)-(5)):		
Search report has been prepared by the EPO or JPO .... $ 930		
International preliminary examination fee paid to USPTO (37 CFR 1.482)...... $ 720		
No international preliminary examination fee paid to USPTO (37 CFR 1.482) but international search fee paid to USPTO (37 CFR 1.445(a)(2))............. $ 790		
Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO................. $1070		
International preliminary examination fee paid to USPTO (37 CFR 1.482) and all claims satisfied provisions of PCT Article 33(2) to (4)............. $ 98		
ENTER APPROPRIATE BASIC FEE AMOUNT	$ 720.00	
Surcharge of $130 for furnishing the oath or declaration later than □ 20 □ 30 mos from the earliest claimed priority date (37 CFR 1.492(e)).	$ 00.00	

CLAIMS	NUMBER FILED	NUMBER EXTRA	RATE	CALCULATIONS	PTO USE ONLY
TOTAL CLAIMS	48 - 20	28	x $ 22	$ 616.00	
INDEPENDENT CLAIMS	12 - 3	9	x $ 82	$ 738.00	
MULTIPLE DEPENDENT CLAIM(S) (if applicable)			+ $270	$ 270.00	

	CALCULATIONS	PTO USE ONLY
TOTAL OF ABOVE CALCULATIONS	$ 2344.00	
Reduction by 1/2 for filing by small entity, if applicable. Verified Small Entity statement must also be filed. (Note 37 CFR 1.9, 1.27, 1.28.)	$ 00.00	
SUBTOTAL	$ 2344.00	
Processing fee of $130 for furnishing the English Translation later than □ 20 □ 30 mos. from the earliest claimed priority date (37 CFR 1.492(f))	$ 00.00	
TOTAL NATIONAL FEE	$ 2344.00	
Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31).	$ 00.00	
TOTAL FEES ENCLOSED	$ 2344.00	
	Amount to be refunded	
	Charged	

a. ■ A check in the amount of $ 2344.00 to cover the above fees is enclosed.
b. □ Please charge my Deposit Account No. 06-1050 in the amount of $ ________ to cover the above fees. A duplicate copy of this sheet is enclosed.
c. ■ The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. 06-1050. A duplicate copy of this sheet is enclosed.

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been ~~met, a~~ petition to revive (37 CFR 1.137(a) or (b) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:

Janis K. Fraser, Ph.D.
FISH & RICHARDSON P.C.
225 FRANKLIN STREET
BOSTON, MA 02110-2804

Janis Fraser
SIGNATURE

Janis K. Fraser, Ph.D.
NAME

34,819
REGISTRATION NUMBER

PCT Rec'd 19 OCT 1998

ATTORNEY DOCKET NO. 00786/292002

Applicant or Patentee: Samuel I. Miller
Serial or Patent No.: 09/068,804
Filed or Issued: May 14, 1998
For: SALMONELLA SECRETED PROTEINS AND USES THEREOF

VERIFIED STATEMENT (DECLARATION) CLAIMING SMALL ENTITY STATUS
(37 CFR 1.9(f) and 1.27(d)) - NONPROFIT ORGANIZATION

I hereby declare that I am an official empowered to act on behalf of the nonprofit organization identified below:

Name of Organization: The General Hospital Corporation
Address of Organization: 55 Fruit Street, Boston, Massachusetts 02114
Type of Organization: NONPROFIT ORGANIZATION

[ ] UNIVERSITY OR OTHER INSTITUTION OF HIGHER EDUCATION
[X] TAX EXEMPT UNDER INTERNAL REVENUE SERVICE CODE (26 USC 501(a) and 501(c)(3))
[ ] NONPROFIT SCIENTIFIC OR EDUCATIONAL UNDER STATUTE OF STATE OF THE UNITED STATES OF AMERICA
(NAME OF STATE:)
(CITATION OF STATUTE:)
[ ] WOULD QUALIFY AS TAX EXEMPT UNDER INTERNAL REVENUE SERVICE CODE (26 USC 501(a) and 501(c)(3)) IF LOCATED IN THE UNITED STATES OF AMERICA
[ ] WOULD QUALIFY AS NONPROFIT SCIENTIFIC OR EDUCATIONAL UNDER STATUTE OF STATE OF THE UNITED STATES OF AMERICA IF LOCATED IN THE UNITED STATES OF AMERICA
(NAME OF STATE:)
(CITATION OF STATUTE:)

I hereby declare that the nonprofit organization identified above qualifies as a nonprofit organization as defined in 37 CFR 1.9(e) for purposes of paying reduced fees under section 41(a) and (b) of Title 35, United States Code with regard to the invention entitled SALMONELLA SECRETED PROTEINS AND USES THEREOF by inventor(s) Samuel I. Miller described in

[ ] the specification filed herewith.
[X] application serial no. 09/068,804, filed May 14, 1998.
[ ] patent no. , issued .

I hereby declare that rights under contract or law have been conveyed to and remain with the nonprofit organization with regard to the above identified invention.

If the rights held by the nonprofit organization are not exclusive, each individual, concern or organization having rights to the invention is listed below* and no rights to the invention are held by any person, other than the inventor, who could not qualify as a small business concern under 37 CFR 1.9(c) or by any concern which would not qualify as a small business concern under 37 CFR 1.9(d) or a nonprofit organization under 37 CFR 1.9(e).

*NOTE: Separate verified statements are required from each named person, concern or organization having rights to the invention averring to their status as small entities. (37 CFR 1.27)

Full Name: ____________________

Address: ____________________

[ ] INDIVIDUAL [ ] SMALL BUSINESS CONCERN [ ] NONPROFIT ORGANIZATION

I acknowledge the duty to file, in this application or patent, notification of any change in status resulting in loss of entitlement to small entity status when any new rule 53 application is filed or prior to paying, or at the time of paying, the earliest of the issue fee or any maintenance fee due after the date on which status as a small entity is no longer appropriate. (37 CFR 1.28(b))

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application, any patent issuing thereon, or any patent to which this verified statement is directed.

Name: David J. Glass, Ph.D.

Title: Associate Director for Patents, Office of Technology Affairs

Address: Massachusetts General Hospital, 13th Street, Bldg. 149, Suite 1101, Charlestown, MA 02129

Signature: [signature] Date: 8/14/98

DAVID J. GLASS, Ph.D.
ASSOCIATE DIRECTOR FOR PATENTS
OFFICE OF TECHNOLOGY AFFAIRS

DAVID J. GLASS
OFFICE OF TECHNOLOGY AFFAIRS
MASSACHUSETTS GENERAL HOSPITAL
13th ST., BLDG. 149, SUITE 1101
CHARLESTOWN, MA 02129

09/068804
Rec'd PCT/PTO 14 MAY 1998

PATENT
ATTORNEY DOCKET NO. 00786/292002

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant : Samuel I. Miller
Serial No.:
Filed : 14 May 1998
Title : SALMONELLA SECRETED PROTEINS AND USES THEREOF

Art Unit:
Examiner:

Int'l Appln. No.: PCT/US96/18504
Int'l Filing Date: 14 November 1996

Box PCT

Assistant Commissioner for Patents
Washington, DC 20231

PRELIMINARY AMENDMENT

Prior to examination, please amend the application as follows:

In the Specification:

Page 1, before the first line, please insert the following paragraph: --This is a continuation of International Patent Application No. PCT/US96/18504, with an international filing date of 14 November 1996, and claims the benefit of prior US provisional application 60/006,733, filed 14 November 1995.--

In the Claims:

In claim 31, line 2, change "40" to --30--.

In the Drawings:

Please substitute the enclosed formal drawings, pages 1/24 to 24/24, for the informal drawings, pages 1/27 to 27/27, now in the application.

"EXPRESS MAIL" Mailing Label Number EM529138070US

Date of Deposit 14 May 1998
I hereby certify under 37 CFR 1.10 that this correspondence is being deposited with the United States Postal Service as **"Express Mail Post Office To Addressee"** with sufficient postage on the date indicated above and is addressed to the Assistant Commissioner for Patents, Washington, D.C. 20231.

Lisa Gray

Lisa G Gray

REMARKS

The specification is amended to recite claimed benefit under 35 USC 120 from International Patent Application PCT/US96/18504. The informal drawings have been replaced with formal drawings. The dependency of claim 31 has been corrected. No new matter has been added by the above amendments.

Please apply any charges not covered, or any credits, to Deposit Account No. 06-1050.

Respectfully submitted,

Date: May 14, 1998

Janis K. Fraser, Ph.D., J.D.
Reg. No. 34,819

Fish & Richardson P.C.
225 Franklin Street
Boston, MA 02110-2804

Telephone: 617/542-5070
Facsimile: 617/542-8906
306961.B11

SALMONELLA SECRETED PROTEINS AND USES THEREOF

Statement as to Federally Sponsored Research

This invention was made with Government support under AI34504 and AI30479 awarded by the National Institutes of Health. The Government has certain rights in the invention.

Background of the Invention

The invention relates to virulence factors of *Salmonella typhimurium*.

Salmonella typhimurium (*S. typhimurium*) enter epithelial cells by a process termed bacterial-mediated endocytosis. *S. typhimurium* stimulates these normally nonphagocytic cells to undergo significant cytoskeletal rearrangements that are visualized as localized membrane ruffling adjacent to the bacteria. Bacteria are then internalized via membrane-bound vacuoles formed from the membrane ruffles.

Several *S. typhimurium* loci have been identified that are required for the induction of bacterial-mediated endocytosis (BME) by epithelial cells. Many of these epithelial-cell signaling loci have a similar chromosomal location, clustered within a 40 kb "virulence island" located between 59 and 60 minutes on the *S. typhimurium* chromosome (Mills et al., *Mol. Microbiol.* 15:749-759, 1995). *InvJ* is a *S. tymphimurium* gene which is thought to encode a secreted protein necessary for BME (Collazo et al., *Mol. Microbiol.* 15:25-38, 1995).

Summary of the Invention

The invention features proteins involved in *Salmonella typhimurium* virulence and/or bacterial-mediated endocytosis. The genes encoding these proteins have now been cloned and their corresponding gene products characterized. Accordingly, the invention

features a substantially pure DNA encoding a *Salmonella* secreted protein (Ssp). By the term "*Salmonella* secreted protein" is meant a *Salmonella*-derived protein, the secretion of which is dependent on the expression of PrgH. In preferred embodiments the invention features substantially pure DNA encoding a *Salmonella typhimurium* secreted protein. By *Salmonella typhimurium* secreted protein is meant as *Salmonella typhimurium* derived protein, the secretion of which is dependent on the expression of PrgH.

One aspect of the invention features a substantially pure DNA molecule which includes the SspB gene; preferably, the DNA includes the DNA sequence of SEQ ID NO: 1, or degenerate variants thereof encoding the amino acid sequence of SEQ ID NO: 5. In another aspect the invention features a substantially pure DNA molecule which includes the SspC gene; preferably, the DNA includes the DNA sequence of SEQ ID NO: 2, or degenerate variants thereof encoding the amino acid sequence of SEQ ID NO: 6. In another aspect the invention features a substantially pure DNA molecule which includes the SspD gene; preferably, the DNA includes the DNA sequence of SEQ ID NO: 3, or degenerate variants thereof encoding the amino acid sequence of SEQ ID NO: 7. In another aspect the invention features a substantially pure DNA molecule which included the SspA gene; preferably, the DNA includes the DNA sequence of SEQ ID NO: 4, or degenerate variants thereof encoding the amino acid sequence of SEQ ID NO: 8. The invention also features a substantially pure DNA molecule which includes the SspB, SspC, SspD, and SspA genes; preferably, the DNA includes the DNA sequence of SEQ ID NO: 15. The invention also features a substantially pure DNA molecule which includes the SspH gene; preferably, the DNA includes the DNA sequence of SEQ ID NO: 13, or degenerate variants thereof encoding

the amino acid sequence of SEQ ID NO: 14. The invention also features a substantially pure DNA molecule which includes the *Salmonella* tyrosine phosphatase A (stpA) gene; preferably, the DNA includes the DNA sequence of SEQ ID NO: 10, or degenerate variants thereof encoding the amino acid sequence of SEQ ID NO:12.

The invention also features a cell into which has been introduced substantially pure DNA encoding an Ssp (or a mutant variant thereof). The substantially pure DNA can be introduced as a portion of a plasmid or other autonomously replicating molecule. In addition the substantially pure DNA can be introduced by homologous recombination. Preferably, the bacterial cell is a *Salmonella* cell; more preferably the bacterial cell is a *Salmonella typhimurium* cell. Cells into which have been introduced substantially pure DNA encoding an Ssp (or mutant variant thereof) can be used as a source of purified Ssp.

The invention includes a substantially pure SspC polypeptide, e.g., a polypeptide which includes an amino acid sequence substantially identical to the amino acid sequence of SEQ ID NO: 6 or an active fragment thereof and a substantially pure SspD polypeptide, e.g., a polypeptide which includes an amino acid sequence substantially identical to the amino acid sequence of SEQ ID NO: 7 or an active fragment thereof. The invention includes a substantially pure SspB polypeptide, e.g., a polypeptide which includes an amino acid sequence substantially identical to the amino acid sequence of SEQ ID NO: 5 (incomplete protein sequence) or an active fragment thereof and a substantially pure SspA polypeptide, e.g., a polypeptide which includes an amino acid sequence substantially identical to the amino acid sequence of SEQ ID NO: 8 (incomplete protein sequence) or an active fragment thereof. The invention includes a

substantially pure full-length SspB polypeptide, e.g., a polypeptide which includes an amino acid sequence substantially identical to the amino acid sequence of SEQ ID NO: 5 (incomplete protein sequence) and the remainder of the SspB sequence. Full-length SspA and SspB genes can be isolated by those skilled in the art using the partial DNA sequences disclosed herein. The invention also includes a substantially pure full-length SspA polypeptide, e.g., a polypeptide which includes an amino acid sequence substantially identical to the amino acid sequence of SEQ ID NO: 8 (incomplete protein sequence) and the remainder of the SspA sequence. The invention also features An active fragment of an Ssp B polypeptide or an SspC polypeptide or an SspD polypeptide is defined as an SspB, SspC, or an SspD polypeptide, respectively, at least 50 amino acids, preferably at least 25 amino acids, more preferably at least 10 amino acids in length having the ability to induce BME in the absence of the full-length version of the corresponding protein. In other preferred embodiments the SspB, SspC, SspD or SspA polypeptide is able to translocate into an epithelial cell, preferably a human epithelial cell. Translocation can be assayed using any suitable assay, e.g., the assay of Sogy et al. (*Molecular Microbiol*. 14:583:94, 1994).

The invention also includes a substantially pure SspH polypeptide, e.g., a polypeptide which includes an amino acid sequence substantially identical to the amino acid sequence of SEQ ID NO:14, or a biologically active fragment thereof.

The invention also includes a substantially pure IagB polypeptide, e.g., a polypeptide which includes an amino acid sequence substantially identical to the amino acid sequence of SEQ ID NO:11, or a biologically active fragment thereof.

Also within the invention is an antibody which binds to a Ssp, e.g., a polyclonal or monoclonal antibody which specifically binds to an epitope of Ssp. Polyclonal and monoclonal antibodies produced against the polypeptides of the invention can be used as diagnostic or therapeutic agents. The invention encompasses not only an intact monoclonal antibody, but also an immunologically-active antibody fragment, e.g., a Fab or $(Fab)_2$ fragment; an engineered single chain Fv molecule. In preferred embodiments, the antibody may be linked to a detectable label, e.g. a radioactive label, fluorescent label, paramagnetic label, or colorimetric label.

The invention also includes a method of detecting a *Salmonella* infection in a mammal which includes the steps of contacting a biological sample derived from the mammal, e.g., a human patient, with a Ssp-specific antibody and detecting the binding of the antibody to a Ssp in the sample. Antibody binding indicates that the mammal is infected with *Salmonella*. The presence of *Salmonella* in a biological sample may also be detected using a method which includes the steps of contacting the sample with a Ssp-encoding DNA, or the complement thereof, under high stringency conditions and detecting the hybridization of the DNA to nucleic acid in the sample. Hybridization indicates the presence of *Salmonella* in the biological sample. By "high stringency" is meant DNA hybridization and wash conditions characterized by high temperature and low salt concentration, e.g., wash conditions of 65°C at a salt concentration of approximately 0.1 × SSC. For example, high stringency conditions may include hybridization at about 42°C in the presence of about 50% formamide; a first wash at about 65°C with about 2 × SSC containing 1% SDS; followed by a second wash at about 65°C with about 0.1 × SSC.

The invention also features a method for detecting the presence of antibodies to an Ssp using all or part of an Ssp protein. The method includes contacting a biological sample with the Ssp protein and measuring the binding of the Ssp protein to an antibody present in the sample.

The invention also features a method of targeting an antigen to an epithelial cell in a mammal which includes the steps of linking the antigen to an Ssp, e.g., SspC or SspD, or active fragment thereof, to produce a Ssp chimeric antigen and administering the chimeric antigen to the mammal.

A method of inducing a cytotoxic T cell immune response in a mammal is also within the invention. This method includes the steps of linking the antigen to an Ssp or active fragment thereof to produce a Ssp chimeric antigen and contacting an antigen-presenting cell, e.g., a Class I major histocompatibility complex (MHC) antigen-expressing cell, with the chimeric antigen.

The invention also features a vaccine which includes a bacterial cell, the virulence of which is attenuated by decreased secretion of a Ssp, and a method of vaccinating an mammal, e.g., a human patient, against a *Salmonella* infection by administering such a vaccine. Preferably, the bacterial cell is a *Salmonella typhimurium* cell, e.g., a *Salmonella enteriditis* cell, or a *Salmonella typhi* cell. A live *Salmonella* cell in which a gene encoding a heterologous antigen is inserted into a Ssp-encoding gene is also included in the invention.

Also within the invention is a substantially pure StpA polypeptide and a method of dephosphorylating a protein which includes the steps of contacting the protein, e.g., a protein at least one tyrosine of which is phosphorylated, with a StpA polypeptide or an active fragment thereof. An active fragment of StpA is defined

as a *Salmonella*-derived polypeptide at least 10 amino acids in length which is capable of removing a phosphate group from a tyrosine residue.

The invention feature live *Salmonella* (particularly *Salmonella typhimurium*) vaccines in which one or more gene required for BME is mutated so as reduce their activity. Among the genes which can be mutated are *SspB*, *SspC*, and *SspD*. Although SspA appears not to be required for BME, it may be useful to mutate this gene as well (preferably in combination with mutation of one or more of the other Ssp genes). Any mutation of these genes which decreases function, including complete or partial deletion and one or more point mutations may be useful. In addition, function of Ssp gene may be impaired by altering its control region.

The invention provides a *Salmonella* vaccine which does not cause transient bacteremia. In general, the invention features a bacterial cell, preferably a *Salmonella* cell, e.g., a *S. typhi*, *S. enteritidis typhimurium*, or *S. cholerae-suis* cell, the virulence of which is attenuated by a first mutation in an Ssp gene. The preferred mutations are loss of function mutations. However, functions causing partial loss of function may be useful if virulence is adequately reduced. Such a bacterial cell can be used as a vaccine to immunize a mammal against salmonellosis.

The *Salmonella* cell may be of any serotype, e.g., *S. typhimurium*, *S. paratyphi A*, *S. paratyphi B*, *S. paratyphi C*, *S. pylorum*, *S. dublin*, *S. heidelberg*, *S. newport*, *S. minnesota*, *S. infantis*, *S. virchow*, or *S. panama*.

The first mutation may be a non-revertible null mutation in one or more of the following genes: *SspB, SspC, or SspD* . Preferably, the mutation is a deletion of at least 100 nucleotides; more preferably, the

mutation is a deletion of at least 500 nucleotides; even more preferably, the mutation is a deletion of at least 750 nucleotides. Mutations in the *prgH* gene or the *prgH* operon can be used for the same purpose.

In preferred embodiments loss or function (partial or complete) is due to decreased expression as a result of a change or mutation, e.g., a deletion, (preferably a non-revertible mutation) at the promoter or other regulatory element of *SspB*, *SspC*, or *SspD* (or some combination thereof).

In another aspect, the invention features a vaccine including a bacterial cell which is attenuated by decrease of expression of a Ssp virulence gene.

The invention also features a live *Salmonella* cell, or a substantially purified preparation thereof, e.g., a *S. typhi*, *S. enteriditis typhimurium*, or *S. cholerae-suis* cell, in which there is inserted into a virulence gene, e.g., an *Ssp* gene, a gene encoding a heterologous protein, or a regulatory element thereof.

In another aspect the invention includes a method of vaccinating an animal, e.g., a mammal, e.g., a human, against a disease caused by a bacterium, e.g., *Salmonella*, including administering a vaccine of the invention.

By "vaccine" is meant a preparation including materials that evoke a desired biological response, e.g., an immune response, in combination with a suitable carrier. The vaccine may include live organism, in which case it is usually administered orally, or killed organisms or components thereof, in which case it is usually administered parenterally. The cells used for the vaccine of the invention are preferably alive and thus capable of colonizing the intestines of the inoculated animal.

By "mutation" is meant any change (in comparison with the appropriate parental strain) in the DNA sequence of an organism. These changes can arise e.g., spontaneously, by chemical, energy e.g., X-ray, or other forms of mutagenesis, by genetic engineering, or as a result of mating or other forms of exchange of genetic information. Mutations include e.g., base changes, deletions, insertions, inversions, translocations or duplications.

A mutation attenuates virulence if, as a result of the mutation, the level of virulence of the mutant cell is decreased in comparison with the level in a cell of the parental strain, as measured by (a) a significant (e.g., at least 50%) decrease in virulence in the mutant strain compared to the parental strain, or (b) a significant (e.g., at least 50%) decrease in the amount of the polypeptide identified as the virulence factor in the mutant strain compared to the parental strain.

A non-revertible mutation, as used herein, is a mutation which cannot revert by a single base pair change, e.g., deletion or insertion mutations and mutations that include more than one lesion, e.g., a mutation composed of two separate point mutations.

Heterologous protein, as used herein, is a protein that in wild type, is not expressed or is expressed from a different chromosomal site, e.g., a heterologous protein is one encoded by a gene that has been inserted into a second gene.

A substantially purified preparation of a bacterial cell is a preparation of cells wherein contaminating cells without the desired mutant genotype constitute less than 10%, preferably less than 1%, and more preferably less than 0.1% of the total number of cells in the preparation.

A substantially pure DNA, as used herein, refers to a nucleic acid sequence, segment, or fragment, which has been purified from the sequences which flank it in a naturally occurring state, e.g., a DNA which has been removed from the sequences which are normally adjacent to the fragment, e.g., the sequences adjacent to the fragment in the genome in which it naturally occurs. The term also applies to DNA which has been substantially purified from other components which naturally accompany the DNA, e.g., DNA which has been purified from proteins which naturally accompany it in a cell.

Other features and advantages of the invention will be apparent from the following description of the preferred embodiments and from the claims.

By "polypeptide" is meant any chain of amino acids, regardless of length or post-translational modification (e.g., glycosylation or phosphorylation).

By "substantially identical" is meant a polypeptide or nucleic acid exhibiting at least 50%, preferably 85%, more preferably 90%, and most preferably 95% sequence identity to a reference amino acid or nucleic acid sequence. For polypeptides, the length of comparison sequences will generally be at least 10 amino acids, preferably at least 20 amino acids, more preferably at least 25 amino acids, and most preferably 35 amino acids. For nucleic acids, the length of comparison sequences will generally be at least 50 nucleotides, preferably at least 60 nucleotides, more preferably at least 75 nucleotides, and most preferably 100 nucleotides.

Sequence identity is typically measured using sequence analysis software (e.g., Sequence analysis software package of the genetics computer group, university of Wisconsin biotechnology center, 1710 university avenue, Madison, WI 53705). Such software

matches similar sequences by assigning degrees of homology to various substitutions, deletions, substitutions, and other modifications. Conservative substitutions typically include substitutions within the following groups: glycine, alanine; valine, isoleucine, leucine; aspartic acid, glutamic acid, asparagine, glutamine; serine, threonine; lysine, arginine; and phenylalanine, tyrosine.

By a "substantially pure polypeptide" is meant a Ssp polypeptide which has been separated from components which naturally accompany it. Typically, the polypeptide is substantially pure when it is at least 60% Ssp by weight. Preferably, the preparation is at least 75%, more preferably at least 90%, and most preferably at least 99%, by weight, Ssp polypeptide. A substantially pure Ssp polypeptide may be obtained, for example, by extraction from a natural source (e.g., *Salmonella* bacterium); by expression of a recombinant nucleic acid encoding a Ssp polypeptide; or by chemically synthesizing the protein. Purity can be measured by any appropriate method, e.g., using column chromatography, polyacrylamide gel electrophoresis, or by HPLC analysis.

A protein is substantially free of naturally associated components when it is separated from those contaminants which accompany it in its natural state. Thus, a protein which is chemically synthesized or produced in a cellular system different from the cell from which it naturally originates will be substantially free from its naturally associated components. Accordingly, substantially pure polypeptides include those derived from one type of prokaryotic organism, e.g., *S. typhimurium*, but synthesized in *E. coli* or another prokaryotic organism.

By "substantially pure DNA" is meant DNA that is free of the genes which, in the naturally-occurring

genome of the organism from which the DNA of the invention is derived, flank the gene. The term therefore includes, for example, a recombinant DNA which is incorporated into a vector; into an autonomously replicating plasmid or virus; or into the genomic DNA of a prokaryote or eukaryote; or which exists as a separate molecule (e.g., a cDNA or a genomic or cDNA fragment produced by PCR or restriction endonuclease digestion) independent of other sequences. It also includes a recombinant DNA which is part of a hybrid gene encoding additional polypeptide sequence, e.g, a hybrid gene encoding a chimeric antigen.

Other features and advantages of the invention will be apparent from the following description of the preferred embodiments thereof, and from the claims.

<u>Detailed Description</u>

Fig. 1 is a diagram of the a genetic map of the 59-60 min region of the *S. typhimurium* chromosome and partial physical map of the restriction endonuclease sites of the *prgH* chromosomal region within the *hil* locus and related plasmids. The horizontal arrows indicate the direction of transcription of the *orf1*, *prgHIJK*, and *org* genes and of the neomycin promoter of the Tn5B50 insertions within the *hil* locus. The vertical arrows indicate and the location of the predicted start of transcription of the *prgHIJK* operon (small arrow) and the location of the two Tn5B50 insertions that define the *hil* locus (large arrows). The open triangle indicates the location of the *prgH1*::Tn*pho*A insertion. Restriction endonuclease sites are as follows: B, *Bam*HI; E, *Eco*RI; H, *Hind*III; S, *Sac*I; Ss, *Ssp*I; V, *Eco*RV; X, *Xho*I.

Fig. 2 is a photograph of a Northern blot assay in which the prgHIJK and *org* transcripts are identified. Blot hybridization of a *prgH* (A), *prgI-J* (B) *prgK* (C), *org* (D), and *pagC* (E) DNA probe to RNA purified from

wild-type (Wt) and phoP constitutive (P^c) *S. typhimurium* strains were grown aerobically to 0.5 optical density units. The bars indicate the RNA markers and are 9488, 6255, 3911, 2800, 1898, and 872 nucleotides (NT) in size from top to bottom.

Fig. 3 is a photograph of a primer extension analysis of RNA isolated from wild-type and $PhoP^c$ *S. typhimurium* strains by using an oligonucleotide primer IB08 corresponding to nucleotides 1217 to 1199 of the *prgH* sequence. Lanes labeled "AGCT" represent dideoxy DNA sequencing reactions. The lane labeled "wt" represents the products of a primer extension reaction initiated with primer IB08 and wild-type RNA as a template, and the lane labeled "P^c" represents the products of a primer extension reaction initiated with the same primer and $PhoP^c$ RNA as a template. Reverse transcription of wild-type RNA with primer IB08 resulted in an approximately 270-nucleotide product corresponding to a predicted transcriptional start at nucleotide 949 of the *prgH* sequence. Abbreviations: wt, wild type strain 14028s; P^c, $PhoP^c$ strain CS022.

Fig. 4A is a diagram showing the similarity and alignment of *prgI, mxiH,* and *yscF* predicted gene products.

Fig. 4B is a diagram showing the similarity and alignment of *prgJ* and *mxiI* predicted gene products.

Fig. 4C is a diagram showing the similarity and alignment of *prgK, mxiJ,* and *yscJ* predicted gene products. For Figs. 4A-4C, residues conserved among each of the predicted gene products are indicated with a plus (+); residues conserved among the *prgI* and either the *mxiH* or *yscF* predicted gene products and between the *prgK* and either the *mxiJ* or *yscJ* predicted gene products are indicated with an asterisk (*). The location of the lipoprotein processing sites (Leu-Xaa-Gly-Cys) of the

prgK, mxiJ, and *yscJ* predicted gene products are indicated by underlining. Predicted protein sequences were compared using the GCG BLAST network service and ALIGN program (Feng et al., *J. Mol. Evol.* 35:351-360, 1987; Higgins et al., *CABIOS* 5:151-153, 1989).

Fig. 5 is a photograph of a SDS-PAGE gel. *Salmonella* proteins found in the culture supernatant of stationary-phase *S. typhimurium* 14028s were compared to proteins isolated from lysed whole cells or cellular fractions (membranes or intracellular soluble proteins). TCA precipitable material from 2 ml of supernatant from cultures of OD_{600} = 2.2 was used. The whole cell, membrane, and soluble lanes contained material from 0.10 ml, 0.35 ml, and 0.15 ml of cells, respectively. Proteins were fractionated in a 12% polyacrylamide gel by SDS-PAGE and stained with Coomassie Brilliant Blue R-250. The molecular masses of protein standards are indicated on the side of the gel as kDa.

Fig. 6 is a photograph of a SDS-PAGE gel showing a comparison of culture supernatant proteins from *S. typhimurium* 14028s and culture supernatants from mutants which are defective in eucaryotic signaling. TCA precipitable material from 2 ml of bacterial culture supernatant was isolated at different times following inoculation: mid-log, OD_{600} = 0.6; late-log / early-stationary, OD_{600} = 1.1; stationary, OD_{600} = 2.2. Proteins were fractionated in a 12% polyacrylamide gel by SDS-PAGE and stained with Coomassie Brilliant Blue R-250. The molecular masses of protein standards are indicated on the side of the gel as kDa. wt, wild type (14028s); p^{c}, $PhoP^{c}$ (CS022); P^{-}, $PhoP^{-}$ (CS015); Δ*hil* (CS451), deleted for the *hil* locus.

Fig. 7 is a photograph of a SDS-PAGE gel showing an analysis of *prgH*::Tn*pho*A and complementation of the insertion mutation by pWKSH5. TCA precipitable material

from 2 ml of supernatant from stationary phase cultures was fractionated in a 10% polyacrylamide gel by SDS-PAGE. Protein was stained with Coomassie Brilliant Blue R-250. The molecular masses of protein standards are indicated on the side of the gel as kDa. wt, wild-type (14028s); IB040, *prgH1*::Tn*pho*A; IB043, *prgH1*::Tn*pho*A with plasmid pWKSH5 containing a 5.1 kb insert of *S. typhimurium* DNA including *prgHIJK*. Supernatant protein bands complemented by pWKSH5 are indicated by arrows (87 kDa and 65 kDa) and a bracket (three bands in the 35-40 kDa range).

Fig. 8 is a photograph of a SDS-PAGE gel showing *Salmonella* secreted proteins (Ssp) concentrated from supernatants of different strains. Each lane contains Ssp collected from 2 ml of culture supernatant. Lanes 1: wild-type *S. typhimurium* SL1344; 2: EE638 (*lacZY11-6*); 3: EE633 (*lacZY4*); 4: VB122 (*hilA*::*kan-112*); 5: EE637 (*invF*::*lacZY11-5*); 6: IB040 (*prgH1*::Tn*pho*A) St: molecular weight standard. Sizes of protein bands are given in kDa. * marks a protein band which was variably present in different preparations of Ssp from the same strains.

Fig. 9 is a diagram showing the chromosomal organization of the *sspBCDA* genes and phenotypes of mutants *sspC*::*lacZY4* (EE633) and *sspA*::*lacZY11-6* (EE638). The chromosomal location of *ssp* with respect to *spaT* and *prgH* is shown. An asterisk (*) indicates partially sequenced genes. Restriction sites in parentheses have only been mapped in the left region of the 11 kb *Eco*RI fragment. Abbreviations of restriction sites are: E: *Eco*RI, B: *Bam*HI, P: *Pvu*II, N: *Nco*I. Invasion of epithelial cells by different *S. typhimurium* strains is given as the percentage of the bacterial inoculum surviving gentamicin treatment. Values represent means and standard errors of the means of three independent experiments, each performed in triplicate. Presence or

absence of *Salmonella* secreted proteins SspA, SspC and SspD in culture supernatants of different strains is indicated by + or -, respectively. The molecular weights in kDa of these Ssp are shown in parentheses.

Fig. 10 is a diagram showing a complementation analysis of EE638. Complementing fragments of chromosomal DNA in a low-copy plasmid are shown according to the chromosomal map. Designations of the plasmids are given in brackets on the left. The positions of the *lac* promoter (P_{lac}) are indicated. Δ indicates a deletion.

Fig. 11 is a photograph of an immunoblot analysis of various strains for expression and secretion of Ssp87. Total cellular proteins from bacteria collected from 0.2 ml of cultures were loaded in lanes designated "C", supernatant proteins from 0.2 ml bacterial culture supernatants were loaded in lanes designated "S". 1: wild type *S. typhimurium*; 2: CS022 ($PhoP^c$); 3: IB040 (*prgH1*::Tn*pho*A); 4: CS451 (Δ*hil*::Tn5-428); 5: EE638 (*sspC*::*lacZY11-6*); 6: EE633 (*sspA*::*lacZY4*).

Fig. 12 is a diagram showing a comparison of the deduced partial amino acid sequence of SspB with the *S. flexneri* homologue IpaB. Bars indicate identical residues, dots indicate gaps introduced in order to maximize similarity according to the GAP program of the GCG package.

Fig. 13 is a diagram showing a comparison of the deduced amino acid sequences of SspC with the *S. flexneri* homologues IpaC. Bars indicate identical residues, dots indicate gaps introduced in order to maximize similarity according to the GAP program of the GCG package.

Fig. 14 is a diagram showing a comparison of the deduced amino acid sequences of SspD with the *S. flexneri* homologues IpaD. Bars indicate identical residues, dots indicate gaps introduced in order to maximize similarity according to the GAP program of the GCG package.

Fig. 15 is a diagram of the amino-terminal sequence derived from the 5'-region of *sspA*. Amino acids determined by amino-terminal sequencing of SspC and SspA are underlined.

Fig. 16 is a photograph of a SDS-PAGE gel showing total soluble Ssp collected from 2 ml of culture supernatants of wild type *S. typhimurium* SL1344 and EE638 (*sspC::lacZY11-6*) transformed with various plasmids. Lanes 1: SL1344 [pWSK29]; 2: EE638 [pWSK29]; 3: EE638 [pCH004 (*sspC*)]; 4: EE638 [pCH005 (*sspCD*)]; 5: EE638 [pCH006 (*sspD*)]; 6: EE638 [pCH002 (*sspCDA*)]; 7: SL1344 [pCH002 (*sspCDA*)]. Lanes 8 and 9 contain soluble Ssp from SL1344 [pWSK29] and EE638 [pWSK29], respectively. The sizes of the protein bands are given in kDa. An asterisk (*) indicates a protein band which was variably present in different preparations of Ssp from the same strains.

Fig. 17 is a photograph of an SDS-PAGE gel showing insoluble Ssp precipitates collected from 2 ml of culture supernatants of wild type *S. typhimurium* SL1344 and EE638 (*sspC::lacZY11-6*) transformed with various plasmids. Lanes 1: SL1344 [pWSK29]; 2: EE638 [pWSK29]; 3: EE638 [pCH004 (*sspC*)]; 4: EE638 [pCH005 (*sspCD*)]; 5: EE638 [pCH006 (*sspD*)]; 6: EE638 [pCH002 (*sspCDA*)]; 7: SL1344 [pCH002 (*sspCDA*)]. Lanes 8 and 9 contain soluble Ssp from SL1344 [pWSK29] and EE638 [pWSK29], respectively. The sizes of the protein bands are given in kDa. An asterisk (*) indicates a protein band which was variably present in different preparations of Ssp from the same strains.

Fig. 18 is a diagram showing the genetic organization of the invasion gene clusters from *S. typhimurium* and *S. flexneri*. The relative positions of each gene are indication and the directions of gene transcription are indicated by arrows. Arrows are not

drawn to scale. Gene clusters conserved in sequence and gene order are indicated by stippling (*inv-spa/mxi-spa*), crosshatching (*prglJK/mxiHI*]), and dark arrows (*ssp/ipa*). Genes with no homologues within the respective regions are shown as open arrows.

Fig. 19 is a depiction of the nucleic acid sequence of *SspB* (missing part of the 5′ end) (SEQ ID NO: 1).

Fig. 20 is a depiction of the nucleic acid sequence of *SspC* (SEQ ID NO: 2).

Fig. 21 is a depiction of the nucleic acid sequence of *SspD* (SEQ ID NO: 3).

Fig. 22 is a depiction of the nucleic acid sequence of *SspB* (missing part of the 3′ end) (SEQ ID NO: 4) and the predicted amino acid sequence SspB (partial c-terminal) (SEQ ID NO: 5).

Fig. 23 is a depiction of the predicted amino acid sequences of SspC (SEQ ID NO: 6), SspD (SEQ ID NO: 7), and SspA (partial animo terminal) (SEQ ID NO: 8).

Fig. 24 is a depiction of the nucleic acid sequences of *iagB* (SEQ ID NO: 9) and *stpA* (SEQ ID NO: 10).

Fig. 25 is a depiction of the predicted amino acid sequences of iagB (SEQ ID NO: 11) and stpA (SEQ ID NO: 12).

Fig. 26 is a depiction of the nucleic acid sequence of prgH (SEQ ID NO: 13).

Fig. 27 is a depiction of the predicted amino acid sequences of prgB (SEQ ID NO: 14).

Fig. 28 is a depiction of the nucleic acid sequence of *SspBCDA* (truncated at 3′ and 5′ ends) (SEQ ID NO: 15).

Fig. 29 is a depiction of the nucleic acid sequence of *prgH* and 5′ and 3′ flanking sequences (SEQ ID NO: 16).

Ssp Proteins and Genes

The Salmonella secreted proteins (Ssp) of the invention have a variety of uses. For example, they can be used as diagnostic reagents, therapeutic agents, and research products. The genes encoding Ssp also have a variety of uses. For example, they can be used as diagnostic reagents. They can also be used to create vaccines including live attenuated vaccines.

Because *Salmonella* infection is a significant health problem and because Ssp proteins are soluble proteins that are found on the surface of *Salmonella*, various Ssp, DNA encoding various Ssp, and antibodies directed against various Ssp are useful in diagnostic assays. Because Ssp are required for optimal virulence, DNA encoding a mutant Ssp having decreased function can be used to create strains of *Salmonella* with reduced virulence. Such strains are useful as live vaccines.

An Ssp (or a portion thereof which can gain entry into the cytoplasm) can be used to translocate a second molecule, e.g., a polypeptide, into the cytoplasm of a cell. This approach can be useful for the induction or priming of cytotoxic lymphocytes (CTL) directed against the second molecule. An Ssp (or a portion thereof capable of translocating an attached second molecule) can be used to introduce a second molecule into the cell cytoplasm for the purpose of drug delivery. Often the second molecule is a polypeptide which is covalently linked to an Ssp (or a portion thereof), e.g., by a peptide bond. Such molecules can be readily produced first preparing a chimeric gene encoding the Ssp (or portion thereof) and the second molecule as a single polypeptide chain. This gene can be used to prepare the fusion protein for administration to a patient. Alternatively, the chimeric gene can be introduced into a

strain of Salmonella which can then be used as either a live vaccine or drug delivery system.

Ssp as Diagnostic Reagents

An Ssp can be used as a diagnostic tool for the detection of *Salmonella* infection in a patient or to evaluate status of an immune response to *Salmonella*. For example, one or more Ssp can be used an antigen in an ELISA assay to detect the presence of *Salmonella*-specific antibodies in a bodily fluid, e.g., blood or plasma, obtained from an infected patient or an individual suspected of being infected with *Salmonella*. Ssp can also be used to test immune cell activation, e.g., T or B cell proliferation or cytokine production, in a sample of patient-derived cells, e.g., peripheral blood mononuclear cells, to detect the presence of a cellular immune response to *Salmonella*.

Polynucleic acids (e.g., primers and probes) encoding all or part of an Ssp can be used in hybridization assays to detect the presence *Salmonella* infection, e.g., using a PCR assay or other probe or primer based assay designed to detect particular DNA sequences.

Antibodies capable of selectively binding a particular Ssp can be used to detect the presence of *Salmonella* in a biological sample. Such antibodies can be produced using standard methods.

Therapeutic Applications of Ssp Fusion Proteins

Fusion proteins comprising all or part of an Ssp and a second protein or polypeptide are useful for a variety of therapeutic applications such as vaccines (e.g., recombinant *Salmonella* vaccines or vaccines against heterologous pathogens), cell targeting agents for delivery of drugs (e.g., cytotoxic agents), and adjuvants, (e.g., to boost an immune response to a co-administered antigen).

To produce a recombinant *Salmonella* vaccine, a gene encoding an Ssp fusion protein can be introduced into a *Salmonella* vaccine. Because Ssp are involved in bacterial mediated endocytosis, the Ssp fusion protein will cause the second polypeptide or protein to be internalized by epithelial cells (or other cells to which the Ssp binds) of the individual to which the vaccine is administered. This internalization can trigger a Type I MHC-mediated response to the second protein or polypeptide. The induction of this response will lead to the induction of CTL (or the priming of CTL) specific for the second protein or polypeptide. The induction or priming of antigen-specific CTL can provide therapeutic or prophylactic benefits.

Purified fusion proteins can be used as recombinant vaccines. Proteins fused to Ssp are specifically targeted to epithelial cells or other cell types to which the Ssp bind; the fusion proteins are then internalized by the targeted cells. Thus, Ssp fusion proteins are useful to generate an immune response to the antigen to which the Ssp is linked or to deliver a therapeutic compound, e.g., a toxin for the treatment of cancer or autoimmune diseases in which the killing of specific cells, i.e., the cells to which a Ssp binds, is desired. Delivery of a toxin linked to a SspC or SspD polypeptide is especially useful in cancer therapy because man types of cancers are of epithelial cell origin.

Ssp fusion proteins which contain all or part of a Ssp linked to a heterologous protein can be made using methods known in the art. Two or more polypeptides may be linked together via a covalent or non-covalent bond, or both. Non-covalent interactions can be ionic, hydrophobic, or hydrophilic.

A covalent linkage may take the form of a disulfide bond. For example, the DNA encoding one of the polypeptides can be engineered to contain a unique cysteine codon. The second polypeptide can be derivatized with a sulfhydryl group reactive with the cysteine of the first component. Alternatively, a sulfhydryl group, either by itself or as part of a cysteine residue, can be introduced using solid phase polypeptide techniques.

A number of other covalent crosslinking agents, e.g., photoreactive crosslinkers, water-soluble crosslinkers, which are commercially available may be used to join a heterologous polypeptide to a Ssp to create a fusion protein. If the fusion protein is produced by expression of fused genes, a peptide bond serves as the link between the components of the fusion protein. Such fusion proteins are produced by expression of a chimeric gene in which sequences encoding all or part of an Ssp are in frame with sequences encoding the second protein or polypeptide. In some circumstances it may be useful to include a linker polypeptide between the Ssp and second protein of polypeptide.

Internalization of the fusion protein may not require the presence of a complete Ssp protein. A internalization-competent portion of an Ssp will be adequate in many circumstances. Whether a particular portion of a selected Ssp is sufficient for internalization can be tested as follows. The selected portion of an Ssp is fused to a calmodulin-dependent adenylate cyclase. If this test fusion protein ii internalized, it will be exposed to calmodulin and the cylcase will be activated. The presence of adenylate cyclase activity can then be used as a measure of internalization. This general approach is described by Sorg et al. (*Molecular Mcrobiol.* 14:583-94, 1994).

Ssp are virulence factors that alter the ability of bacteria to be internalized by specific populations of host cells and to induce an immune response. *Salmonella* with mutations in genes encoding Ssp are useful in the manufacture of live *Salmonella* vaccines with altered cell tropism.

Deletion or overexpression of Ssp in *Salmonella* can be used to target strains or fusion proteins to various mammalian cell types. Invasion of epithelial cells or macrophages can be selected depending on the Ssp mutated. For example, use of *Salmonella* as an antigen or drug delivery vehicle can be optimized by deleting part or all of a gene encoding a Ssp involved in bacterial mediated endocytosis (or mutating such a gene to impair Ssp function), thereby minimizing the ability of *Salmonella* to invade epithelial cells (and therefor maximizing antigen delivery to antigen presenting cells such as macrophages). In this manner, strains with mutated Ssp genes can be used to modulate the host immune system. Deletion of Ssp genes in *Salmonella* can also be used to alter the ability of *Salmonella* to stimulate IL-8 secretion by epithelial cells.

Fusions of antigens to Ssp genes can be used to facilitate an immune response to the linked antigens for the purpose of generating an antigen-specific cytotoxic T cell response in a patient. For example, Ssp fusions to viral antigens are useful as therapeutic vaccines for diseases such as AIDS and *Herpes genitalis* in which the generation of a cytotoxic T cell (CTL) response is desired. Delivery of antigens in this manner favors the generation of an antigen-specific CTL response because the Ssp portion of Ssp fusion protein mediates translocation of the fusion protein across eucaryotic cell membranes into the intracellular compartments in the cytoplasm of cells which participate class I MHC-mediated

antigen processing and presentation, i.e., the generation of class I MHC-restricted antigen-specific CTLs.

Fusion proteins which include all or part of a Ssp linked to a cytotoxic molecule can be used to target a cytotoxic molecule to a specific cell type, e.g., an epithelial cell-derived cancer cell, which would then by killed by the cytotoxic agent. Cytotoxic fusion proteins can be synthetically or recombinantly produced and administered directly to a patient. Alternatively, live *Salmonella* expressing a cytotoxic Ssp fusion protein can be administered and allowed to produce and secrete the fusion protein *in vivo*.

Ssp are also useful as adjuvants to boost the immunogenicity of antigens with which they are delivered or to which they are chemically or recombinantly linked. Ssp that have enzymatic effects, e.g., phosphatase activity, on certain types of eucaryotic cells can be used to promote specific types of immune responses such as TH2 or TH1 T cell responses. Since these proteins are secreted and are likely taken up in the cytoplasm of eucaryotic cells, gene fusions to these proteins are likely to be more immunogenic and more efficient in inducing the development of an immune response, particularly a class I MHC-restricted CTL response.

Various oral and parenteral delivery systems are known in the art and can be used to deliver the Ssp polypeptides and/or chimeric antigens of the invention, such as encapsulation in liposomes, or controlled release devices. The compositions of the invention can be formulated in a pharmaceutical excipient in the range of approximately 10 μg/kg and 10 mg/kg body weight..

The compositions and methods of the invention provide the tools with which to construct better vaccines against *Salmonella* infection and for the prevention and treatment of other diseases, e.g., cancer and AIDS, by

using *Salmonella* secreted proteins as carriers of heterologous antigens, e.g., tumor antigens or viral antigens, either as purified components or as hybrid proteins produced in live *Salmonella* vaccine strains.

Ssp and Attenuated Bacterial Strains

Deletion or mutation of one or more Ssp genes can be used to attenuate vaccine strains. For instance deletion of Ssp genes leads to lack of neutrophil transmigration across epithelial cell monolayers (a model system that correlates well with the ability of certain strains to cause gastroenteritis).

Vaccine strains are usually administered at doses of 1×10^5 to 1×10^{10} cfu/single oral dose. Those skilled in the art can determine the correct dosage using standard techniques.

Research products

Ssp with enzymatic activity, e.g., *Salmonella* tyrosine phosphatase (stpA), can be used as reagents for protein modification. StpA catalyzes the release of phosphate groups from tyrosine residues in proteins, and thus, is especially useful in the field of signal transduction. Since a number of eucaryotic and procaryotic signal transduction proteins are regulated by the phosphorylation and dephosphorylation of tyrosine residues, stp can be used to deactivate or activate these proteins, thereby altering intracellular signal transduction. Thus, Stp can be used as a research tool to study and evaluate phosphorylation-regulated signal transduction pathways.

Modification of Ssp and Ssp Variants

When an Ssp is being used to translocate a second molecule into a eukaryotic cell, it may be useful increase expression of the Ssp (or Ssp fusion protein) so that BME is increased. Increased expression of *sspC*, *sspD* and other *ssp* genes may be accomplished using

methods known in the art, e.g., by introducing multiple copies of the gene(s) into the bacterial cell or cloning the Ssp-encoding DNA under the control of a strong promoter.

Under other circumstances it may be desirable to increase uptake of a bacterial strain, e.g., a *Salmonella* strain, by a macrophage in a mammal by impairing the normal invasion mechanism of the strain. This can be accomplished by decreasing expression of the DNA encoding the SspC and/or SspD (and thereby decreasing secretion of Ssp and/or SspD polypeptides) and administering the cell to the mammal. Ssp expression may be reduced using methods known in the art, e.g., insertion of a transposon (Tn) into the gene, deletion of some or all of the gene, mutating a gene upon which SspC and/or SspD expression depends, e.g., *prgH*, e.g., a deletion or Tn insertion in the *prgHIJK* operon. Instead of decreasing the expression of *sspC* and/or *sspD*, the method may include the step of impairing the function of one or both of the gene products, e.g., by Tn insertion, deletion mutagenesis, or by impairing the secretory pathway by which the gene products are secreted such that the gene products are produced but not effectively transported to the extracellular environment.

Example 1: PhoP/PhoQ Transcriptional Repression of *S. typhimurium* Invasion Genes: Evidence for a Role in Protein Secretion

The PhoP-repressed *prgH* locus of *S. typhimurium* may be important for signaling epithelial cells to endocytose *S. typhimurium*. The following series of experiments demonstrate that the *prgH* locus is an operon of four genes encoding polypeptides of 392 amino acids (prgH), 80 amino acids (prgI), 101 amino acids (prgJ), and 252 amino acids (prgK). These experiments also

demonstrate that expression of the 2.6-kb *prgHIJK* transcript is reduced when PhoP/PhoQ is activated, suggesting that PhoP/PhoQ regulates *prgHIJK* by transcriptional repression. Further, analysis of the culture supernatants from wild-type *S. typhimurium* revealed the presence of at least 25 polypeptides larger than 14 kDa. Additional experiments demonstrated that *prgH1*::Tn*pho*A, *phoP* constitutive (PhoP^c), and *hil* deletion mutants have significantly defective supernatant protein profiles. A further set of experiments described below demonstrate that both the invasion and supernatant protein profile defects of the *prgH1*::Tn*pho*A mutant can be complemented by a 5.1 kb plasmid that included *prgHIJK*. Taken together these results suggest that PhoP/PhoQ regulates extracellular transport of proteins by transcriptional repression of secretion determinants and that secreted proteins are likely involved in signaling epithelial cells to endocytose bacteria.

The following reagents and procedures were used to evaluate the *prgH* locus.

<u>Bacterial Strains, Growth and Conditions</u>

S. typhimurium strain ATCC 14028s (American Type Culture Collection, Bethesda, MD) is a virulent wild-type parent strain from which all other *Salmonella* strains described in Example 1 were derived. Bacterial strains and plasmids are described in Table 1. Luria-Bertani broth (LB) was used as rich bacterial growth medium. Antibiotics were added to LB broth or agar in the following concentrations: ampicillin, 25 μg/ml; chloramphenicol, 50 μg/ml; kanamycin, 45 μg/ml.

<u>DNA sequencing and analysis</u>

Double-strand templates were sequenced by the dideoxy-chain termination method known in the art as modified for use with Sequenase^{TM} (US Biochemicals, Corp.) and [α-^{35}S]dATP. Computer analysis of the DNA sequence

was accomplished with the GENEPRO (Riverside Scientific, Riverside, CA) and Wisconsin package (GCG, version 7) programs. The nucleotide sequence of the *prgHIJK* locus has been deposited in GeneBank under accession number U21676.

RNA extraction, RNA blot analyses, and primer extension analyses

RNA was isolated from mid-log phase cultures (OD_{600} = 0.5) of aerobically-grown (with shaking) and microaerophically-grown (without shaking) *Salmonella* strains using a standard hot phenol procedure (Pulkkinen et al., *J. Bacteriol.* 173:86-93, 1993). For RNA blots, 20 μg of RNA was diluted in H_2O and incubated for 15 minutes at 55°C in 50% formamide, 17.5% formaldehyde in 1 × Northern buffer (0.36 M Na_2HPO_4-$7H_2O$, 0.04 M NaH_2PO_4-H_2O). Samples were run on 1% agarose gels containing 6% formaldehyde and 1 × Northern buffer and were transferred to Gene Screen Plus membranes (NEN/Dupont). RNA was crosslinked to the membrane using a Stratalinker™ UV crosslinker (Stratagene, La Jolla, CA). Membranes were hybridized and washed according to the manufacturer's protocol.

The DNA probes for RNA-DNA and DNA-DNA blot hybridization were obtained from recombinant plasmid DNA by restriction endonuclease digestion or by polymerase chain reaction (PCR) using the GeneAmp™ PCR kit (Perkin-Elmer/Cetus). The following DNA probes were synthesized: a 841-bp *prgH* probe from the oligonucleotide primers IB07 (5'-CCAGGTGGATACGGA-3'; SEQ ID NO: 17; nucleotides 1198 to 1212) and IB19 (5'-TAGCGTCCTCCCCATGTGCG-3'; SEQ ID NO: 18; nucleotides 2039 to 2021); a 433-bp *prgI-prgJ* probe from the primers IB26 (5'-CCGGCGCTACTGGCGGCG-3'; SEQ ID NO: 19), nucleotides 2304 to 2321) and DP04 (5'AGCGTTTCAACAGCCCCG-3'; SEQ ID NO: 20), nucleotides

2737 to 2719); a 341-bp *prgK* probe from primers DP03 (5'-CGGGGCTGTTGAAACGC-3'; SEQ ID NO: 21), nucleotides 2720 to 2736) and DP08 (5'-AACCTGGCCTTTTCAG-3'; SEQ ID NO: 22), nucleotides 3060 to 3045); a 724-bp *org* probe from primers DP15 (5'-GGCAGGGAGCCTTGCTTGG-3'; SEQ ID NO: 23), nucleotides 3774 to 3792) and DP17 (5'-GTGCCTGGCCAGTTCTCCA-3'; SEQ ID NO: 24); and a 608-bp *pagC* probe from a *Psi* and StuI restriction-endonuclease digest of pWPL4 that contains the wild-type *pagC* gene. DNA probes were radiolabelled using a standard method of random priming with [α-^{32}P]dCTP.

For primer extension analyses, oligonucleotide primers (0.2 picomoles) were end-labelled with [γ-^{32}P]dATP (NEN/Dupont), annealed to *S. typhimurium* RNA (20μg) and extended with reverse transcriptase (Gibco BRL, St. Louis, MO). Reactions were electrophoresed in 6% polyacrylamide, 8 M urea gels adjacent to sequencing reactions initiated with primers used for cDNA synthesis.

<u>DNA blot hybridization analysis</u>

Chromosomal DNA was isolated, restriction endonuclease digested, size fractionated in agarose gels, and transferred to GeneScreen Plus membranes (NEN/Dupont). For dot blot hybridization experiments, high stringency hybridization was performed according to standard methods at 65°C using radiolabelled probes.

<u>Protein isolation and analysis</u>

Bacteria were grown in LB, with shaking at 37°C. Bacterial cultures were chilled to 4°C and centrifuged at 154,000 × g for 1.7 hours. The supernatant was carefully removed and trichloroacetic acid (TCA) was added to a final concentration of 10%. The precipitates were collected by centrifugation at 69,000 × g for 1 hour, rinsed with cold acetone, dried and stored at 4°C. The bacterial cell pellet was fractionated to obtain periplasmic, cytoplasmic, and membrane fractions.

Samples were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) on a 10-12% polyacrylamide (0.1 M Tris pH 8.45, 0.1% SDS) gel using a standard Tris-glycine buffer system or standard Tris-tricine buffer system. TCA precipitates were mixed with sample buffer (250 mM Tris pH 6.8, 2% SDS, 0.0025% bromophenol blue, 5.0% β-mercaptoethanol, 10% glycerol) and heated to 100°C for 5 minutes. Proteins were visualized by staining with Coomassie Brilliant Blue R-250.

<u>Enzyme assays</u>

Presence of the marker enzymes, alkaline phosphatase (periplasm) and β-galactosidase (cytoplasm) were used to assess fraction purity. A plasmid, pPOS3, containing an arabinose-inducible *phoA* gene, was inserted into wild-type strain 14028s by transformation and moved into other strains using P22 bacteriophage-mediated transduction. Addition of arabinose (0.02%) to the culture medium induced transcription of the *phoA* gene. Determination of alkaline phosphatase activity of strains containing pPOS3 was performed using the substrate p-Nitrophenyl phosphate according to standard methods. The results were expressed in standard units for β-galactosidase (Miller, J.H., 1972, Experiments in Molecular Genetics, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., pp. 352-355). β-galactosidase was produced from a strain with a *mudJ*-generated gene fusion of *msg* and *lacZ*. The gene, *msg*, is constitutively expressed and not PhoP regulated. β-galactosidase activity of strains carrying *msg*::*MudJ* was measured using routine methods (Miller et al., *supra*).

Table 1. Bacterial strains, plasmids and relevant properties

S. typhimurium	Relevant genotype
ATCC 14028s	Wild Type
CS002	*pho-24*
CS019	*phoN2zxx::6251Tn10d-Cm*
IB040	CS019 with *prgH1::TnphoA*
IB043	IB040 with pWKSH5
CS015	*phoP-102::Tn10d-Cm*
CS451	14028s derivative of EE451 with Δ*hil*
Escherichia coli	
DH5α	F⁻Φ8Φd*lacZ*ΔM15Δ(*lacZYA-argF*)U169*end*A1 *rec*A1*hsd*R17*deo*R*thi*-1*sup*E44λ*gyr*A96*rel*A1
Plasmids	
pIB01:	pUC19 (amp^R) containing a 10.7-kb *Eco*RV fragment with *prgH1*::Tn*phoA* (kan^R)
pVB3	pUC19 containing a 5.9-kb *Hind*III-*Eco*RI fragment of the *prgH* locus
pWKSH5	pWKS30 (amp^R) containing a 5.1-kb *Hind*III fragment of *prgH* locus
pWPL4	pUC19 containing a 5.0-kb *Eco*RV fragment of the *pagC* locus
pPOS5	pBR322 containing arabinose-inducible PhoA

Cloning and sequencing of *prgH*

The DNA containing the *prgH1*::Tn*phoA* gene fusion was cloned based upon information derived from the physical map of restriction endonuclease sites surrounding the transposon insertion (Fig. 1) (Behlau et al., *J. Bacteriol.* 175:4475-4484, 1993, hereby incorporated by reference). Chromosomal DNA from strain IB040 containing the *prgH1*::Tn*pho*A insertion was digested with the restriction endonuclease *Eco*RV and ligated into SmaI digested pUC19 to generate a library of recombinant plasmids. These recombinant plasmids were transformed into *Escherichia coli* (*E. coli*) DH5α. A recombinant plasmid containing a 10.7 kb *EcoRV* fragment was identified by selecting for kanamycin resistance (Tn*pho*A

encoded) and was designated pIB01 (Fig. 1). DNA hybridization analysis of strain IB040 with a radiolabelled 1.5-kb *Hind*III-*Sac*I-generated DNA fragment of pIB01 resulted in hybridization to an approximately 10.7-kb *Eco*RV DNA fragment. This was approximately 7.7 kb (the size of Tn*pho*A) larger than the 3-kb fragment present in the wild-type strain ATCC 14028s. This probe also hybridized strongly to plasmid pVB3 that contained the 5.9 kb *Hind*III-*Eco*RI fragment of the *hil* locus (Fig. 1), confirming the location of the *prgH* locus within this region. This data indicated DNA containing the *prgH1*::Tn*pho*A insertion had been cloned.

The DNA sequence of the 4,034-bp *Hind*III-*Ssp*I fragment (within which the Tn*pho*A insertion in *prgH* was localized) was determined by sequencing plasmid pIB01 containing the cloned *prgH1*::Tn*pho*A allele. This sequence was confirmed by DNA sequencing of pWKSH5 containing the wild-type *prgH* allele (Fig. 1). Information from DNA sequence of the *prgH1*::*phoA* fusion junction was used to determine the direction of transcription and correct reading frame of *prgH*. Tn*pho*A was inserted after nucleotide 1548 within an open reading frame that extended from nucleotides 981 to 2156. *prgH* was predicted to encode a 392 amino acid polypeptide with a calculated M_r of 44,459 daltons and pI of 5.86. The N-terminal portion of *prgH* was found to have a stretch of nonpolar residues followed by the motif Leu-Xaa-Gly-Cys at residues 24 to 27 (corresponding to nucleotides 1050 to 1061) characteristic of the processing site of bacterial lipoproteins. There was a strong hydrophobic domain (amino-acid residue 144 to 154, corresponding to nucleotides 1410 to 1433) upstream of the Tn*pho*A insertion.

Analysis of the nucleotide sequence located upstream of *prgH* revealed an additional open reading

frame from nucleotides 665 to 222, termed *orf1*, likely to be oppositely transcribed from *prgH*. The intergenic region between *orf1* and *prgH* was 216 nucleotides. *orf1* was predicted to encode a gene product of 148-amino-acid residues with a calculated M_r of 17,186. The start codon of *orf1* was preceded by a potential ribosome binding site at 7 to 11 nucleotides 5′ to the predicted start of translation (5′-AAAGG-3′, nucleotides 676 to 672) suggesting that this open reading frame was translated. The *orf1* predicted gene product had no signal sequence nor any strong hydrophobic domains.

<u>Identification of *prgI*, *prgJ*, and *prgK*</u>

Analysis of the nucleotide sequence located downstream from *prgH* revealed four additional open reading frames that were predicted to be transcribed in the same direction and form an operon: (a) nucleotides 2184 to 2423; (b) nucleotides 2445 to 2747; (c) nucleotides 2747 to 3502; and (d) nucleotide 3476 to beyond the 3′ *SspI* site. The first three of these four open reading frames identified were designated *prgI*, *prgJ*, and *prgK* respectively. *prgI*, *prgJ*, and *prgK* were predicted to encode gene products of 80 amino acids (M_r, 8865 daltons), 101 amino acids (M_r, 10,929 daltons), and 252 amino acids (M_r, 28,210 daltons). The predicted gene products encoded by *prgI* and *prgJ* did not contain a signal sequence or strong hydrophobic domains. The predicted gene product encoded by *prgK* contained a N-terminal hydrophobic region followed by a potential lipoprotein processing site from amino-acid residue 15 to 18 (corresponding to nucleotides 2788 to 2800). The fourth open reading frame corresponded in DNA sequence to the *S. typhimurium* oxygen-regulated gene (*org*).

<u>*prgH-K* transcription is negatively regulated by PhoP/PhoQ</u>

To determine whether *prgH* was negatively regulated by PhoP/PhoQ, RNA isolated from wild-type (ATCC 14028s)

and $PhoP^{C}$ (CS022) strains of *S. typhimurium* were analyzed. In numerous RNA blot analyses, the prgH-specific DNA probe hybridized with an approximately a 2600-nucleotide RNA from the wild-type strain (Fig. 2). The size of the RNA that hybridized to the prgH probe was similar to that of the *prgH-K* open reading frame predicted from the DNA sequence (i.e., 2600 vs. 2522 nucleotides). In contrast, no transcript was seen when equal amounts and similar quality of RNA (as assessed by methylene blue staining) isolated from the $PhoP^{C}$ strain was probed with prgH-specific DNA (Fig. 3). In comparison, when the same RNA preparations were hybridized with a *pagC*-specific probe, an approximately 1100-nucleotide *pagC* transcript was highly expressed in the $PhoP^{C}$ strain (Fig. 2), consistent with the constitutive phenotype of *pag* gene expression in the $PhoP^{C}$ mutant (Pulkkinen et al., *J. Bacteriol.* 173:86-93, 1991, hereby incorporated by reference). These results indicate that regulation of *prgH* occurs at the level of transcription.

Primer extension analysis was performed to obtain information on the possible initiation site of prgH transcription. Based on this analysis, the start of prgH transcription was predicted to begin approximately 32 nucleotides upstream from the *prgH* translational start (Fig. 3). Several different primers were used that resulted in primer extension products of differing lengths, but all predicted that transcription initiated at this site. The predicted -10 (5'-TAATCT-3') and -35 (5'-TTCATC-3') regions are similar to the consensus sequences for typical α70 E. *coli* promoters. Similar to the results of RNA blot hybridization analysis, a primer extension product was detected only with RNA isolated from wild-type *S. typhimurium* and not with RNA isolated from the $PhoP^{C}$ strain (Fig. 3).

The size of the RNA that hybridized to the prgH-specific probe suggested that *prgH-K* could form a transcriptional unit. Therefore, to determine whether *prgI-K* formed an operon that was regulated by PhoP/PhoQ, RNA blot hybridization and primer extension analysis were performed using DNA probes and primers specific to the *prgI*, *prgJ*, and *prgK* open reading frames. Similar to the results with *prgH,* the *prgI-J-* and *prgK*-specific DNA probes hybridized with an approximately 2600-nucleotide RNA isolated from wild-type *S. typhimurium* and not with RNA from the $PhoP^C$ strain (Fig. 2). No primer extension products less than 350 nucleotides were detected using RNA isolated from either the wild-type or $PhoP^C$ strains using *prgI*, *prgJ*, and *prgK* primers. These primers were from 1662 to 2332 nucleotides downstream from the predicted start of *prgH* transcription. These findings indicated that *prgH-K* were transcribed as an operon, heretofore referred to as *prgHIJK*. Furthermore, this operon was likely to be transcribed from the prgH promoter and was negatively regulated by PhoP at the level of transcription.

<u>org is not regulated by PhoP/PhoQ</u>

Although the above results suggested that the *prgHIJK* transcriptional unit did not include org, experiments were performed to test this possibility. Blot hybridization analysis was performed with RNA isolated from wild-type *S. typhimurium* and an *org*-specific DNA probe. As shown in Fig. 2, two distinct transcripts hybridized to the *org* probe: an approximately 1400-nucleotide abundant RNA and a minor RNA of approximately 3800 nucleotides. The size of the smaller RNA was similar to that of the *org* open reading frame (1400 vs. 1236 nucleotides). In comparison, only the major 1400-nucleotide RNA was seen when RNA from the $PhoP^C$ strain was hybridized with the org-specific DNA

probe, suggesting that the 3800-nucleotide RNA was PhoP repressed.

A minor RNA of approximately 3800 nucleotides also was detected in long exposure of wild-type RNA blots that were hybridized with either the prgH, *prgI-J*, or *prgK* probes, suggesting possible cotranscription of *prgHIJK* and *org*. However, both the major (1400 nucleotide) and minor (3800 nucleotide) transcripts were detected when RNA isolated from the *prgH1*::Tn*pho*A strain was hybridized with the *org* probe, indicating that the *prgH1*::Tn*pho*A insertion was not polar on either of the *org* transcripts. Because expression of an *org*::*lacZY* fusion was shown to be increased approximately fourteen fold in low-oxygen compared with high-oxygen tension, RNA from wild-type and $PhoP^C$ strains that were grown aerobically or microaerophically to an optical density at 260 nm of 0.5 were compared by blot hybridization with the *org*-specific DNA probe. No substantial difference was seen in the relative amounts of RNA transcripts detected in wild-type or $PhoP^C$ strains grown under these conditions. These data indicate that *org* did not form part of the *prgHIJK* operon and that it was not regulated by PhoP/PhoQ.

<u>The prgI, prgJ, and prgK predicted polypeptides are similar to *S. flexneri* Mxi and *Y. enterocolitica* Ysc proteins</u>

The sequences of the five predicted polypeptides (PrgH, PrgI, PrgJ, PrgK, and Orf1 were compared with the protein sequences translated from the GeneBank library using BLAST network software. This comparison revealed similarity between the predicted products of *prgI*, *prgJ*, and *prgK* and the MxiH, MxiI, and MxiJ proteins of *S. flexneri* . Each of the these polypeptide sequences were similar over their entire length, with 65% (PrgI vs. MxiH), 38% (PrgJ vs. MxiI), and 46% (PrgK vs. MxiJ) of

positions occupied by identical residues (Figs. 4A-4C). The *prgI* and *prgK* predicted gene products were also similar to the YscF and YscJ proteins, respectively, of *Y. enterocolitica*, with 28% and 308 of positions occupied by identical residues. The Poisson probabilities were highly significant for each of these comparisons. No protein similar to the *prgH* or *orf1* predicted polypeptides was detected in the protein sequence library.

<u>Isolation of proteins from *S. typhimurium* culture supernatants</u>

The role of *prgHIJK* in *S. typhimurium* protein secretion was analyzed by examination of the proteins present in cell culture supernatant. Culture media of wild-type bacteria was collected for protein analysis by centrifuging stationary phase cultures at 154,000 × g for 1.7 hours. From 6-8 μg/ml of protein was precipitated by addition of trichloroacetic acid (TCA) to overnight culture supernatants. The TCA-precipitable material in 2 ml of supernatant then was fractionated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (Fig. 5). Approximately 25 protein bands, ranging in molecular mass from 18-87 kDa, were detected by Coomassie brilliant blue staining.

To rule out the possibility that the supernatant protein bands represented proteins released from lysed cells, soluble and membrane fractions of whole cells and whole cell lysates were compared with proteins from the supernatant by SDS-PAGE (Fig. 5). Many of the major polypeptides in the supernatant (e.g., the polypeptide with molecular mass of 87 kDa) were not the major proteins in the other cellular fractions. Conversely, major intracellular soluble proteins and membrane proteins (e.g., the 36 kDa OmpC porin) were not detected in the supernatant in this analysis. In addition,

following centrifugation, the overnight culture media from bacteria expressing alkaline phosphatase (a periplasmic protein) and β-galactosidase (a cytoplasmic protein) always contained less than 9% and 1%, respectively, of the whole-cell activity of these enzymes. Although some of the supernatant protein bands may represent degradation products of larger protein species, these data indicate that *S. typhimurium* was capable of significant protein secretion.

To determine the timing of release of polypeptides in to the supernatant and to test for an effect of PhoP regulon mutations on secretion, supernatants from CS015, with a null mutation in PhoP ($PhoP^-$), CS022 ($PhoP^C$), and wild-type bacteria (ATCC 14208s) were compared. As shown in Fig. 6, the quantity of protein increased for each strain when supernatant samples taken from mid-log-phase (OD_{600} = 0.6), late-log/early-stationary-phase (OD_{600} = 1.1), and stationary-phase (OD_{600} = 2.2) were compared. However, the pattern of major protein bands detected for each strain was unchanged from mid-log to stationary phase (Fig. 6).

<u>Altered supernatant protein profiles of mutants defective in signaling epithelial cells</u>

At each phase of growth examined, a similar pattern and quantity of protein was detected in the culture supernatants of $PhoP^-$ strain CS015 and wild-type bacteria (Fig. 6). In contrast, the protein level of 2 ml of $PhoP^C$ strain CS022 supernatant was 24% of wild type levels. At least 10 major protein bands seen in the wild-type supernatant were greatly reduced or undetectable by Coomassie blue staining of the CS022 supernatant, especially those of higher molecular weight (Fig. 6). In addition, four major protein bands appeared to be increased in amount in CS022 compared with wild-type supernatant (31.5 kDa, 30 kDa, 23 kDa, and 20

kDa) (Fig. 6). Although this result could be due to degradation of higher molecular weight polypeptides, these data suggest that the PhoP^{C} mutant likely was defective in synthesis or secretion of Ssp.

The defect observed with the PhoP^{C} mutant was consistent with *prg* gene products having a role in protein secretion. Therefore, the Ssp of strains having transposon insertion or deletion of *prgHIJK* were compared to wild-type bacteria (ATCC 14028s) by SDS-PAGE. As observed for the PhoP^{C} mutant, IB040 (*prgH1*::Tn*pho*A) and CS451, containing a 10-kb deletion of *hil* locus (Δ*hil*) DNA, each had a pronounced defect in their Ssp profile compared with the wild-type strain (Fig. 6 and 7). IB043 and CS451 culture supernatants contained 100% and 62%, respectively, of wild-type protein levels. At least 5 and 11 major protein bands seen in the wild-type supernatant were greatly reduced or undetectable by Coomassie blue staining of the IB040 and CS451, respectively. Five protein bands [87 kDa, 65 kDa, and three in the 35-40 kDa range (Fig. 7), two of which run as a doublet under different electrophoretic conditions (Fig. 6)] were undetectable in the supernatants of CS022, IB040, and CS451. These findings indicated that the presence of at least some of the products of the *prgHIJK* operon were necessary for synthesis or secretion of Ssp.

The defect in bacterial mediated endocytosis associated with *prgH1*::Tn*pho*A was complemented by a low-copy number plasmid, pWKSH5, containing a 5.1-kb fragment including *prgHIJK*, *org*, and *orf1*. Consistent with this observation, the *prgH1*::Tn*pho*A mutant carrying pWKSH5 (strain IB043) had a supernatant protein profile similar to that of wild type (Fig. 7). Of the five protein bands undetectable or greatly reduced in culture supernatants of *prgH1*::Tn*pho*A, each was detected in IB043 and three of them were increased in amount (87 kDa, 65

kDa, and 35 kDa) compared with wild-type supernatants. This finding demonstrates a correlation between the ability to secrete proteins and induction of epithelial cell bacterial mediated endocytosis.

The *prgH* locus is important for *S. typhimurium* to induce endocytosis by epithelial cells

The defect in BME of the *prgH1*::Tn*pho*A mutant is complemented by a plasmid containing 5.1 kb of DNA from this region, indicating that the gene or genes disrupted by the *prgH1*::Tn*pho*A insertion are important for BME. Analysis of the DNA sequence of this region identified six potential open reading frames that could be affected by this transposon insertion. As depicted in Fig. 1, five of these open reading frames, namely those designated *prgH-K* are either disrupted (i.e., *prgH)* or are 3′ to the *prgH1*::Tn*pho*A insertion. The *orf1* translational start is 884 nucleotides upstream from the TnphoA insertion and that *orf1* is predicted to be oppositely transcribed from the *prgHIJK* operon.

An approximately 2600 nucleotide PhoP-repressed transcript was detected when RNA was hybridized with *prgH-*, *prgI-J-*, or *prgK*-specific DNA probes. In contrast, the predominant transcripts detected with *org* was smaller (approximately 1400 nucleotides), was not altered in the *prgH1*::Tn*pho*A mutant, and was not repressed by PhoP. Primer extension analysis of the potential start site of transcription, the size of the *prgHIJK* transcript, and the presence of a potential transcriptional terminator immediately downstream of *prgK* also were consistent with transcription terminating before *org*.

In addition to the major transcripts of *prgHIJK* and *org*, a minor PhoP-repressed transcript of approximately 3800 nucleotides also was detected in multiple RNA blots hybridized with the *org* and *prgH*,

prgI-J, or *prgK* DNA probes. This minor RNA was similar in size to the combined *prgHIJK* and *org* open reading frames (i.e., 3731 nucleotides) and, thus, could represent cotranscription of *prgHIJK* and *org*. However, both the 3800- and 1400-nucleotide transcripts were detected in RNA from the *prgH1*::Tn*pho*A mutant suggesting that the 3800-nucleotide RNA did not represent cotranscription of *prgHIJK* and *org*. These data indicate that one or more genes in the *prgHIJK* operon are important to BME of epithelial cells.

A PhoP constitutive mutation repressed the synthesis of approximately 20 prg-encoded cell-associated protein species (Miller et al., *J. Bacteriol.* 172:2485-2490, 1990, herein incorporated by reference). Although PhoP/PhoQ has been shown to transcriptionally activate *pag* (Miller et al., *Proc. Natl. Acad. Sci. USA* 86:5054-5058, 1989, herein incorporated by reference; Pulkkinen et al., *supra*, herein incorporated by reference), the mechanism of protein repression by PhoP/PhoQ had not been characterized prior to the present studies. No transcript was detected when RNA from the PhoP constitutive mutant was probed with *prgH*-, *prgI-J*-, or *prgK*-specific DNA, indicating that the *prgHIJK* operon was negatively regulated by PhoP/PhoQ at the level of transcription. Thus, PhoP/PhoQ can both activate and repress transcription of virulence genes.

Consistent with the role of one or more of the products of *prgHIJK* in bacterial mediated endocytosis and possibly in protein secretion, a low-copy plasmid containing 5.1 kb of DNA (IB043), including *prgHIJK*, *org*, and *orf1*, complemented both the bacterial mediated endocytosis defect and the supernatant protein profile defect of the *prgH1*::Tn*pho*A mutant. Based upon its similarity to MxiJ and YscJ, which are membrane-associated lipoproteins that are necessary for

export and secretion of Ipa and Yops protein respectively, the *prgK* gene product is most likely to have such a role in bacterial mediated endocytosis and protein secretion. Similar to PrgK, PrgH was predicted to be a membrane lipoprotein. However, in contrast to *prgI-K*, which are similar to plasmid-encoded genes of *Shigella* and *Yersinia* spp., a *prgH* DNA probe hybridized to chromosomal DNA but not virulence-plasmid DNA from *Shigella* spp.

Neither they nor the *prgI* or *prgJ* predicted gene products have signal sequences or long hydrophobic domains that suggest their cellular localization. However, the location of these genes within operons that encode secretion determinants suggests that they may have a role in this process.

The predicted gene products of the *prgHIJK* operon were found to be similar to gene products required for protein secretion in other bacterial species. An analysis of proteins present in culture supernatants of *S. typhimurium* was performed. These experiments revealed that the supernatants of wild-type cultures contained a large number of protein bands, whereas strains with mutations affecting the *prgH* locus, including *prgH1*::Tn*pho*A, Δ*hil* and PhoP^{C} were each defective in protein secretion as assessed by Ssp profiles. This analysis suggested that PhoP/PhoQ could control protein secretion, at least in part, by repressing *prgHIJK* whose products could form part of a secretion machinery. Furthermore, the finding that PhoP^{C} and Δ*hil* mutants were associated with greater defects in their Ssp profile compared with the *prgH1*::Tn*pho*A mutant suggested that more than one mechanism may be involved in protein secretion and that gene products encoded by the 10 kb region that is deleted in the *hil* mutant also contribute to secretion of Ssp.

Since the strains with altered Ssp profiles were each impaired in signaling epithelial cells, these data suggest that Ssp are involved in signaling such cells to initiate BME. The finding that five Ssp were missing from culture supernatants of the *prgH* mutant suggested that one or more of these proteins were specifically involved in BME, e.g., Ssp and/or *prgHIJK* gene products may form a structure on the surface of *S. typhimurium* which induces bacterial mediated endocytosis.

S. typhimurium strains with transposons inserted between *prgH* and *spa* that result in reduced bacterial mediated endocytosis were also missing a subset of the Ssp missing from the *prgHIJK* mutant. DNA sequence analysis of the regions flanking the transposon insertions revealed deduced protein sequences that were similar to IpaB and IpaD of *S. flexneri*. These data suggest that the transposon insertions define an operon in *S. typhimurium* that encodes Ipa homologues.

Example 2: *Salmonella typhimurium* Secreted Invasion Determinants

Two *Salmonella typhimurium* secreted protein (Ssp) mutants with transposon insertions located between *spaT* and *prgH* were identified. One mutant lacks the 87 kDa Ssp, while the other lacks Ssp of 87, 42, and 36 kDa. The invasiveness of these mutants implicates the 42 and 36 kDa Ssp, but not the 87 kDa Ssp in invasion. DNA sequencing of this region identified two complete and two partial open reading frames (designated *sspB*, *sspC*, *sspD*, and *sspA*).

The deduced amino acid sequences of *sspBCDA* are homologous to *Shigella flexneri* secreted proteins IpaB, IpaC, IpaD, and IpaA. Complementation analyses and amino-terminal sequencing showed that *sspC* and *sspA* encode the 42 kDa and the 87 kDa Ssp and that both

proteins are secreted without amino-terminal processing. SspA is abundantly secreted by wild type bacteria but is completely retained within the cellular fraction of a mutant in *prgHIJK* encoding part of the Ssp secretion apparatus. A precipitate containing SspC and three major Ssp of 63, 59, and 22 kDa was isolated from culture supernatants of wild type bacteria. These data indicate that major secreted invasion determinants of *S. typhimurium* are structurally and functionally homologous to *S. flexneri* Ipa proteins.

The following reagents and experimental procedures were used to characterize Ssp.

Construction of plasmids and strains:

To construct pCH002, pVV8-1 was cut with *Eco*RI, the 11 kb fragment eluted from a 1% agarose gel and cloned into the *Eco*RI site of pWSK29. In pCH002, transcription of *sspCDA* is driven from the *lac* promoter. pCH004 was constructed by cloning the 3 kb *Bam*HI fragment from pCH002 into the *Bam*HI site of pWSK29. pCH005 contains the 4 kb *Eco*RI-*Pvu*II fragment from pCH002 cloned into *Eco*RI-*Hinc*II restricted pWSK29. pCH006 was constructed by restriction of pCH005 with *Nco*I and religation of the 1.7 kb and the 5 kb fragment. The correct orientations of the cloned inserts were confirmed by appropriate restriction analyses.

PCR of a chromosomal fragment of EE638 comprising the 5'-region of Tn5*lacZY* and adjacent DNA was performed in three independent experiments by using primers OL 1 (5' CGCGGATCCATTATGGGATGTATCGG 3'; SEQ ID NO: 25) and OL2 (5' CCGGCAGCAAAATGTTGCAG 3'; SEQ ID NO: 26). The 0.8 kb amplified DNA fragments were then restricted with *BamH*I and cloned into pWSK29 for sequencing. All three sequences were identical.

Strain VB122 (*hil*A::*kan*-112) was constructed as follows: the mutation was originally constructed on a

plasmid by inserting a *kan* cassette (Pharmacia Biotech, Piscataway, NJ) in a *Hinc*II site in the 5′ region of the *hilA* coding sequence. The plasmid-encoded *hilA::kan-112* mutation was recombined into the chromosome, and the chromosomal mutation was confirmed by PCR analysis.

Mutant EE633 (*lacZY*4) was isolated by screening for oxygen regulated gene fusions created by random Tn5*lacZY* insertions in *S. typhimurium* VV114 (*hilA::kan-114*) and further selection for insertions linked to a *hilA::kan-114* by P22 transduction into *S. typhimurium* SL1344 and selection for Tet^R and Kan^R.

<u>Media and growth conditions for bacterial cultures</u>:

Bacteria were grown in LB broth at 37°C. If necessary, selection was carried out using 50 μg/ml ampicillin, 10 μg/ml tetracycline, or 25 μg/ml kanamycin.

<u>Preparation and analysis of *S. typhimurium* supernatant proteins</u>:

Bacterial cultures were grown for 16 to 17 hours in 12 ml LB in 1.5 × 14 cm glass tubes at 37°C on a TC-7 roller (New Brunswick, Edison, NJ) at 50 rev./min. Soluble proteins from culture supernatants were obtained as described above. Precipitates in the culture were retrieved, rinsed 5 times with 1 ml H_2O, dissolved in sample buffer (4% SDS, 12% glycerol, 5% β-mercaptoethanol, 0.05 M Tris-HCl pH 6.8, 0.01% bromphenol blue), and resolved in 10% polyacrylamide gels using SDS-PAGE and a Tris-Tricine buffer.

<u>Immunoblotting</u>:

Whole cell samples were prepared from overnight cultures using standard methods with the additional step of filtering the culture through a Whatman 1 qualitative paper filter (Whatman International, Maidstone, Kent, England) before centrifugation. The proteins were resolved by SDS-PAGE and transferred to nitrocellulose by

electroblotting using a conventional transfer buffer. Western blots were incubated with polyclonal rabbit serum prepared against the 87 kDa Ssp. The immunogen was purified by SDS-PAGE and injected into New Zealand White rabbits (Charles River, Wilmington, MA). Serum was collected after two booster injections and subsequently absorbed with an acetone powder prepared from *S. typhimurium* strain EE63. Horseradish peroxidase-labelled goat anti-rabbit antibodies were used to label the primary antibodies and were visualized using chemiluminescence (ECL, Amersham, International, Buckinghamshire, England)

<u>Invasion assays</u>:

Invasion of HEp-2 epithelial cells was carried out according to the method of Behlau et al. (*J. Bacteriol.* 175:4475-4484, 1993). To minimize epithelial cell detachment from the bottom of the assay wells after bacterial uptake, the following modifications were introduced: invasion time was reduced from 90 to 60 min and gentamicin treatment was performed for 15 min with 100 μg/ml gentamicin, conditions which were shown to kill 99% of a bacterial culture of 2×10^8 cells/ml.

<u>N-terminal protein sequencing</u>:

Proteins separated by SDS-PAGE were blotted on PVDF membranes (Bio-Rad, Hercules, CA) and stained with Ponceau-S. Blotted proteins were sequenced using an ABI 470A protein sequencer with 120A PTH-AA analyzer.

Table 7: Strains and plasmids used in this example

Bacterial strain	Marker
E. coli DH5α	F-Φ80d*lac*ZΔM15Δ(*lac*ZYA-*arg*F)U169*end*A1 *rec*A1*hsd*R17(r_K^-, m_K^+)*deo*R*thi*-1*sup*E44λ*gyr*A96*rel*A1
S. typhimurium SL1344	wild type
VV114	*hil*::*kan*-114
VB122	Kan^R, *hilA*::*kan*-112
EE637	Tet^R, *invF*::*lacZY11*-5
EE633	Tet^R, *sspA*::*lacZY*4
EE638	Tet^R, *sspC*::*lacZY11*-6
S. typhimurium (ATCC14028s)	wild type
CS451	14028sΔ*hil*::Tn5-428
CS022	*pho*-24 (PhoP^c)
IB04	*prg*H1::Tn*pho*A

Plasmid	Marker
pWSK29	Amp^R
pVV8-1	Tet^R
pVV71	Amp^R
pCH002	Amp^R, *ssp*CDA, *hil*A
pCH004	Amp^R, *ssp*C
pCH005	Amp^R, *ssp*CD
pCH006	Amp^R, *ssp*D

Identification of *S. typhimurium* Mutants with Transposon Insertions in Genes Encoding Ssp

To identify genes encoding Ssp, Tn5*lacZY* mutants of *S. typhimurium* SL1344 with transposon insertions located within the 40 kb "virulence island" (59-60 min. of the *S. typhimurium* chromosome) were analyzed for changed patterns in Ssp. An insertion in *invF* (*invF*::*lacZY11-5*), the first gene of the *inv-spa* operon, and a *hilA*::*kan-122* insertion in VB122 led to major defects in the pattern of Ssp similar to a mutation in the *prgHIJK* operon (*prgH1*::Tn*pho*A) which has implicated in *S. typhimurium* protein secretion (see Example 1). Specifically, all three mutants lack 5 major Ssp of 36, 38, 42, 63 and 87 kDa, while the *hilA*::*kan-112* insertion leads to loss of some lower molecular weight protein bands in addition to these 5 Ssp (Fig. 8, lanes 4, 5, 6).

Two other mutants exhibited detectable loss of only one and of three Ssp, respectively. The supernatants from the mutant strain EE633 containing the fusion *lacZY4* were missing a protein of 87 kDa, while supernatants from the mutant strain EE638, containing fusion *lacZY*11-6, were missing protein species of 87, 42 and 36 kDa. In addition, supernatants from EE638 showed an increased abundance of a 63 kDa Ssp (Fig. 8, lanes 2, 3). Tn5*lacZY* in EE638 maps approximately 2.5 kb downstream from *spaT* while in EE633 the transposon maps 5.5 kb downstream from *spaT* as determined by Southern hybridization and PCR analyses. Both transposons were inserted in the same orientation (Fig. 9). A degenerate pool of oligonucleotides synthesized according to the sequence of the 12 amino-terminal amino acids of the 87 kDa protein (VTSVRTQPPVIM; SEQ ID NO: 27), hybridized specifically to a 5.5 kb *Bam*HI fragment in pVV71 which comprises sequences between *hilA* and *spaT* (Fig. 9). These data indicate that the 87 kDa Ssp is encoded in the chromosomal region adjacent to the transposon insertions. Tn5*lacZY* in EE633 is likely to be directly within the gene encoding the 87 kDa Ssp, while Tn5*lacZY* in EE638 is likely to be inserted within one of the genes encoding the 42 and the 36 kDa Ssp having a polar effect on the synthesis of the other two Ssp missing in supernatants of this mutant.

Secretion of the 87 Ssp kDa Ssp is Dependent on *prgHIJK*

Since it was possible that the absence of the 87 kDa Ssp (Ssp87) in supernatants of EE633 and EE638 was due to impaired secretion rather than expression, whole cell lysates and supernatants of various strains were analyzed by immunoblotting with antiserum raised against partially purified Ssp87. Fig. 11 shows that Ssp87 of wild type *S. typhimurium* is found mainly in the supernatant, although some of the protein is detected in

the cellular fraction (Fig. 11, lane 1). In contrast to wild type bacteria, all of the protein is found in the cellular fraction of the *prgH1*::Tn*pho*A mutant IB040 (lane 3). Ssp87 could not be detected in the cellular fractions nor in supernatants of various invasion and secretion mutants: CS022 (PhoP^C), a mutant which constitutively represses PhoP regulated genes (Miller et al., *J. Bacteriol.* 172:2485-2490, 1990, hereby incorporated by reference) (lane 2), CS451 (Δ*hil*::Tn5-428) carrying a 10 kb chromosomal deletion of the *hil* locus between 59 and 60 min. (lane 4), EE638 (*lacZY*11-6) (lane 5), and EE633 (*lacZY*4) (lane 6). The signal at 51 kDa in the supernatant fraction of wild type bacteria might represent a degradation product of Ssp87, while the faint band at 34 kDa is likely nonspecific hybridization since it is present in all bacteria analyzed. These results demonstrate that lack of Ssp87 in supernatants of EE633 and EE638 is due to impaired expression while lack of Ssp87 in supernatants of IB040 (*prgH1*::Tn*pho*A) is caused by impaired secretion of the protein. These results further show that expression of the gene encoding Ssp87 is affected by the *hil* deletion and that expression of Ssp87, either directly or indirectly, is repressed by PhoP.

<u>Strain EE638, but not EE633, Is Markedly Deficient in Invasion</u>

To determine the function of the 87, 42, and 36 kDa Ssp in invasion of epithelial cells, the ability of strain EE638 and EE633 to invade HEp-2 cells was analyzed. EE638 showed more than a 100-fold reduction in invasiveness when compared to wild type bacteria, while EE633 exhibited invasion levels comparable to wild type bacteria (Fig. 9). These results suggested that the 36 and/or the 42 kDa Ssp but not the 87 kDa Ssp are required for epithelial cell invasion. In addition, observation

of interactions between these mutants and PtK2 cells by time-lapse videomicroscopy indicated that the ability of EE638 to induce epithelial cell membrane ruffling is also markedly reduced, while EE633 induced localized membrane ruffles at a frequency similar to wild type *S. typhimurium*.

<u>The Tn5*lacZY* Insertions in EE638 and EE633 Define a Chromosomal Region Encoding Ssp *S. typhimurium* Homologues of the *Shigella ipaBCDA* Operon</u>

To determine the gene(s) affected by the transposon insertions in EE638 and EE633, part of a 11 kb *Eco*R1 subclone of pVV8-1 was sequenced. Two complete and two partial open reading frames (ORFs), positioned in the same transcriptional direction, were identified (Fig. 9). The deduced gene products of the complete ORFs exhibit similarity to *Shigella* secreted proteins IpaC and IpaD (31% identity, 47% similarity; 37% identity, 56% similarity) respectively, and therefore were designated *sspC* and *sspD* (see Fig. 13 and Fig. 14). The gene products of the complete open reading frames were designated *sspC* and *sspD*. The amino acid sequence derived from the 5'-end of *sspC* was identical to the amino-terminal sequence of the 42 kDa Ssp (underlined in Fig. 13). The deduced gene product of the partial ORF located immediately upstream from *sspC* shows 47% identity (67% similarity) to the carboxyterminal portion of *S. flexneri* secreted protein IpaB and was designated *sspB* (Fig. 12). The ORF starting immediately downstream of *sspD* was designated *sspA*. The amino acid sequence deduced from the 5' end of an ORF starting immediately downstream from *sspD* did not exhibit similarity to IpaA. However, DNA sequencing of internal parts of the gene predicted that the protein encoded by this gene, designated *sspA*, is similar to IpaA. Nevertheless, the sequence of amino acids 2-13 (underlined in Fig. 15) was

identical to the amino-terminal sequence of the 87 kDa Ssp (see above). *sspB*, *sspC*, *sspD*, and *sspA* are separated by 27, 70, and 15 bp, respectively, and putative ribosome binding sites precede *sspC*, *sspD*, and *sspA* .

The amino acid similarities of Ssp to Ipas do not extend over the entire lengths of the proteins. The similarities between SspC/IpaC and SspD/IpaD are highest in the carboxy-terminal regions, while the central parts of SspB and IpaB are conserved (see Fig. 12, 13, and 14). These similarities could reflect conservation in regions of the proteins required for secretion and/or invasion. Although both SspC and SspD appear to be secreted by the same mechanism, no obvious similarities or motifs common to these proteins were detected, thus implying conformational rather than sequential features in the secretion of proteins by type III secretion pathways.

The precise insertion of Tn5*lacZY* in EE638 was determined by cloning and sequencing of a PCR product comprising the 5′ region of the transposon and upstream chromosomal sequences and was shown to be located 189 bp downstream from the ATG start codon of *sspC*. The order of the *ssp* genes and the Ssp profile of EE638 indicate that the transposon insertion in *sspC* is polar on expression of *sspD* and *sspA* and that these genes are likely to be organized in a singly transcribed unit.

Both *sspC* and *sspD* are Necessary for *S. typhimurium* Invasion of Epithelial Cells

A complementation analysis was carried out to determine the minimal fragment necessary for complementation of the epithelial cell invasion defect of EE638 (*sspC*::*lacZY*11-6) as well as for reconstitution of Ssp. All analyzed fragments were cloned downstream from the *lac* promoter in the 6-8 copies/chromosome vector pWSK29. As shown in Fig. 10, a 3.9 kb *Eco*RI-*Pvu*II

fragment comprising *sspC* and *sspD* in pCH005 was sufficient to complement the invasion defect of EE638 to wild type levels. When analyzed for Ssp, EE638 [pCH005] showed a pattern of Ssp similar to the wild type strain [pWSK29] except for the missing 87 kDa protein (SspA) (Fig. 16, lane 4). EE638 transformed with pCH002 carrying an 11 kb *Eco*R1 fragment was partially complemented for invasion as well as for all 3 missing Ssp (Fig. 10 and Fig. 16, lane 6). In contrast, EE638 transformed with plasmids that contained either *sspC* or *sspD* alone (pCH004 and pCH006, respectively) were not complemented for invasion but showed reconstitution of the 42 kDa Ssp (SspC) or the 36 kDa Ssp (SspD), respectively (Fig. 10 and Fig. 16, lanes 3 and 5). In addition, the abundancy of a 63 kDa Ssp, which was found to be more abundant in supernatants of EE638, was reduced in supernatants of strains EE638 [pCH005], EE638 [pCH006], and EE638 [pCH002] and of SL1344 [pCH002]. These results demonstrate that both *SspC* and *SspD* are necessary for invasion of epithelial cells and indicate that *SspC* encodes the 42 kDa Ssp while the 36 kDa Ssp is likely to be encoded by *SspD*. In addition, complementation of the invasion defect of EE638 with pCH005 indicates that invasiveness is not influenced by the observed changes in the abundancy of the 63 kDa Ssp.

<u>A Precipitate Found in *S. typhimurium* Culture Supernatants Contains Highly Abundant SspC and Other Proteins</u>

Supernatants from *S. typhimurium* wild type cultures contained a precipitate that, when solubilized in reducing SDS sample buffer, separates into at least four highly abundant protein bands of 63, 59, 42 and 22 kDa on SDS-PAGE (see Fig. 17, lane 1). Protein precipitates were also found in culture supernatants of EE638 and EE633, but not in supernatants of *S.*

typhimurium mutants with global defects in protein secretion [CS022 (PhoP^{C}), IB040 (*prgH*::Tn*pho*A), CS451 (Δhil::Tn5-428) and VB122 (*hilA*::*kan-112*). *S. typhimurium* 14028s, the wild type parent of CS022 and IB040, showed the same protein pattern of precipitated material as SL1344]. The precipitate from EE633 cultures showed a similar composition to that of wild type precipitate by SDS-PAGE analysis. In contrast, a major protein band of 42 kDa was absent from the precipitate isolated from cultures of EE638 (Fig. 17, lane 2). Amino-terminal sequencing of this 42 kDa Ssp identified it as encoded by *SspC*. The identity of the amino-terminal protein sequence (MLISNVGINPAAYLN; SEQ ID NO: 28) with the amino acid sequence derived from the 5'-region of *SspC* (Fig. 13) shows that no amino-terminal processing of SspC occurs prior to its release into the supernatant.

SDS-PAGE analyses of precipitated material from culture supernatants of EE638 [pCH004 (*SspC*)] and EE638 [pCH005 (*SspCD*)] showed a pattern similar to wild type [pWSK29] material (Fig. 17, lane 3 and 4), confirming that the respective plasmids complemented the mutant for secretion of SspC. Protein patterns of soluble Ssp and precipitates isolated from untransformed cultures of SL1344 or EE638 were identical to those shown in Fig. 16, 17, lane 1 and 2, respectively. Precipitate of EE638 [pCH006 (*SspD*)] was found to be similar to precipitate from EE638 [pWSK29] except for reduced abundancy of a 63 kDa protein band (Fig. 17, lane 5). The precipitate from EE638 [pCH002 (*SspCDA*)] contained an additional major protein band of approximately 51 kDa, which was also present in precipitate from SL1344 [pCH002] (Fig. 17, lanes 6, 7). Comparison of precipitated proteins to soluble Ssp on SDS-PAGE (Fig. 17, lanes 8, 9) showed that SspC in the precipitate has the same electrophoretic

mobility as the 42 kDa soluble Ssp. These data suggest that the 42 kDa soluble Ssp is identical to precipitated SspC.

SspC and SspA are secreted proteins of 42 and 87 kDa, as demonstrated by amino-terminal sequencing and by complementation analyses. It is further likely that the 36 kDa protein encoded by *SspD* is secreted, since lack of a 36 kDa Ssp in supernatants of EE638 (*lacZY*11-6) was complemented by transformation of this mutant with plasmids containing *SspD*. The 63 kDa Ssp is the protein likely to be encoded by SspB.

SspA, SspB, SspC, and SspD appear to be targets of the *inv-spa-prgHIJK* encoded secretion apparatus, since these Ssp are missing in supernatants of mutants affecting expression or regulation of *inv-spa* and *prgHIJK* (Fig. 8). Typical for proteins secreted by type III secretion pathways, no amino-terminal processing of SspA and SspC was observed. The dependency of Ssp secretion on *prgHIJK* was further proven by demonstrating that SspA is abundantly secreted by wild type cells, while it is completely retained in the cellular fraction of the *prgH1* ::Tn*pho*A mutant IB040 (Fig. 11). The 38 kDa Ssp of the five major Ssp dependent on the *inv-spa-prgHIJK* secretion apparatus may be the product of the *invJ* invasion locus.

The immunoblot analysis of SspA secretion suggests that expression of the gene encoding SspA is negatively controlled by the virulence two component regulatory system PhoP/PhoQ. PhoP/Q has a global effect on protein secretion which is partially due to negative transcriptional regulation of *prgHIJK* (see Example 1).

The SspBCDA genes are located between the large inv-spa gene cluster and *prgHIJK* at 59 minutes on the *S. typhimurium* chromosome. Fig. 18 shows the relative positions of the invasion genes in *S. typhimurium* in comparison to their *S. flexneri* homologues, which are

clustered in a 31 kb region of a large virulence plasmid. The invasion genes cluster in three groups (*inv-spa/mxi-spa*, *Ssp/ipa*, and *prgIJK/mxiHIJ*) which exhibit conserved gene structure and organization, suggesting that these genes were acquired by horizontal gene transfer. Acquisition by horizontal gene transfer is further supported by the fact that these *S. typhimurium* invasion genes are within a 40 kb "virulence island" which, despite the otherwise high overall genetic similarity between *S. typhimurium* and *E. coli* K-12, is unique to *S. typhimurium*. However, the three invasion gene clusters from *S. flexneri* and *S. typhimurium* are in different relative positions to each other and are interspersed between non-homologous genes, thus implying multi-recombinational events in the evolution of these genetic regions.

In addition to soluble Ssp the supernatants of *S. typhimurium* cultures contained a flocculent precipitate consisting of SspC and three other major protein species of 63 (Ssp 63), 59 (Ssp 59), and 22 (Ssp 22) kDa. The combination and abundancy of Ssp in the precipitate from *S. typhimurium* cultures is strikingly different from that in the soluble fraction (see Fig. 17). Though Ssp, including SspC, are found in both the precipitate and the soluble fraction, SspD, even when overproduced, was not detected in the precipitate. This emphasizes the difference in composition of precipitate and soluble fraction and supports the possibility of specific protein-protein interactions between the four Ssp leading to precipitate formation.

OTHER EMBODIMENTS

Using reagents derived from partial cDNA clones of an Ssp, e.g., SspA, the isolation of a full-length cDNA encoding the Ssp is well within the skill of those

skilled in the art of molecular biology. For example, a radiolabelled probe made from a known partial cDNA sequence can be used to identify and isolate from a library of recombinant plasmids cDNAs that contain regions with identical to the previously isolated cDNAs. The screening of cDNA libraries with radiolabelled cDNA probes is routine in the art of molecular biology (see Sambrook et al., 1989, *Molecular Cloning: a Laboratory Manual*, second edition., Cold Spring Harbor Press, Cold Spring Harbor, N.Y). The cDNA can be isolated and subcloned into a plasmid vector, and the plasmid DNA purified by standard techniques. The cDNA insert is sequenced using the dideoxy chain termination method well known in the art (Sambrook et al, *supra*). Oligonucleotide primers corresponding to bordering vector regions as well as primers prepared from previously isolated cDNA clones can be employed to progressively determine the sequence of the entire gene.

Similar methods can be used to isolate Ssp which are related to SspA, SspB, SspC, or SspD. To isolate related Ssp, a probe having a sequence derived from (or identical to) all or a portion of SspA, SspB, SspC, or SspD can be used to screen a library of *Salmonella* DNA (or cDNA). DNA encoding a related Ssp will generally hybridize at greater stringincy than DNA encoding other proteins. This approach can be used to identify *Salmonella typhimurium* Ssp as well as Ssp of other *Salmonella*.

<u>Generation of Monoclonal Antibodies</u>:

Monoclonal antibodies can be generated to purified native or recombinant gene products, e.g., Ssp, by standard procedures, e.g., those described in Coligan et al., eds., *Current Protocols in Immunology*, 1992, Greene Publishing Associates and Wiley-Interscience). To generate monoclonal antibodies, a mouse is immunized with

the recombinant protein, and antibody-secreting B cells isolated and immortalized with a non-secretory myeloma cell fusion partner. Hybridomas are then screened for production of specific antibodies and cloned to obtain a homogenous cell population which produces a monoclonal antibody. For example, hybridomas secreting the desired antibodies can be screened by ELISA. Specificities of the monoclonal antibodies can be determined by the use of different protein or peptide antigens in the ELISA. Useful quantities of antibodies can be produced by either the generation of ascites fluid in mice or by large scale *in vitro* culture of the cloned antibody-producing hybridoma cell line. Antibodies can be purified by various chromatographic procedures known in the art, such as affinity chromatography on either immobilized Protein A or Protein G.

The invention also includes DNA encoding other Ssp (e.g., Ssp 54, Ssp 42, and Ssp 22) found in cell supernatants. Those skilled in the art can readily clone the corresonding genes based on the amino terminal sequence or the corresponding protein. The amino terminal sequence of Ssp54 is MNNLTLSXFXKVG (SEQ ID NO: 29). The amino terminal sequence of Ssp42 is MLISNVGINPAAYLN (SEQ ID NO: 30). The amino terminal sequence of Ssp 22 is TKITLSPQNFFI (SEQ ID NO: 31).

CLAIMS

1. Substantially pure DNA encoding a *Salmonella* secreted protein (Ssp).

2. The DNA of claim 1, wherein said DNA comprises the SspB gene.

3. The DNA of claim 2, wherein said DNA comprises the DNA sequence of SEQ ID NO: 1 or degenerate variants thereof encoding the amino acid sequence of SEQ ID NO: 5.

4. The DNA of claim 1, wherein said DNA comprises the SspC gene.

5. The DNA of claim 4, wherein said DNA comprises the DNA sequence of SEQ ID NO: 2 or degenerate variants thereof encoding the amino acid sequence of SEQ ID NO: 6.

6. The DNA of claim 1, wherein said DNA comprises the SspD gene.

7. The DNA of claim 6, wherein said DNA comprises the DNA sequence of SEQ ID NO: 3 or degenerate variants thereof encoding the amino acid sequence of SEQ ID NO: 7.

8. The DNA of claim 1, wherein said DNA comprises the SspA gene.

9. The DNA of claim 8, wherein said DNA comprises the DNA sequence of SEQ ID NO: 4, or degenerate variants thereof encoding the amino acid sequence of SEQ ID NO: 8.

10. The DNA of claim 1, wherein said DNA comprises the SspB gene, the SspC gene, the SspD gene and the SspA gene.

11. The DNA of claim 10, wherein said DNA comprises the DNA sequence of SEQ ID NO: 15.

12. The DNA of claim 1, wherein said DNA comprises the SspH gene.

13. The DNA of claim 12, wherein said DNA comprises the DNA sequence of SEQ ID NO: 13, or degenerate variants thereof encoding the amino acid sequence of SEQ ID NO: 14.

14. The DNA of claim 1, wherein said DNA comprises the stpA gene.

15. The DNA of claim 14, wherein said DNA comprises the DNA sequence of SEQ ID NO: 10 or degenerate variants thereof encoding the amino acid sequence of SEQ ID NO: 12.

16. A cell which contains the DNA of claim 1.

17. A method of inducing uptake of a bacterial cell by an epithelial cell in a mammal, comprising increasing expression of the DNA of claim 4 or 6 in said cell and administering said cell to said mammal.

18. The method of claim 17, wherein said bacterial cell is a *Salmonella* cell.

19. A method of inducing uptake of a bacterial cell by a macrophage in a mammal, comprising decreasing expression of the DNA of claim 4 or 6 and administering said cell to said mammal.

20. A substantially pure SspC polypeptide.

21. The polypeptide of claim 20, comprising an amino acid sequence substantially identical to the amino acid sequence of SEQ ID NO: 6.

22. An active fragment of the polypeptide of claim 21.

23. A substantially pure SspD polypeptide.

24. The polypeptide of claim 23, comprising an amino acid sequence substantially identical to the amino acid sequence of SEQ ID NO: 7.

25. An active fragment of the polypeptide of claim 24.

26. A substantially pure SspH polypeptide.

27. The polypeptide of claim 26, comprising an amino acid sequence substantially identical to the amino acid sequence of SEQ ID NO: 14.

28. An active fragment of the polypeptide of claim 27.

29. A substantially pure IagB polypeptide.

30. The polypeptide of claim 29, comprising an amino acid sequence substantially identical to the amino acid sequence of SEQ ID NO: 11.

31. An active fragment of the polypeptide of claim 40.

32. An antibody which binds to a Ssp.

33. A method of detecting a *Salmonella* infection in a mammal comprising contacting a biological sample derived from said mammal with the antibody of claim 32 and detecting the binding of said antibody to a Ssp in said sample, wherein said binding indicates that said mammal is infected with *Salmonella*.

34. A method of detecting the presence of *Salmonella* in a biological sample comprising contacting said sample with a Ssp-encoding DNA under high stringency conditions and detecting the hybridization of said DNA to nucleic acid in said sample, wherein hybridization indicates the presence of *Salmonella* in said biological sample.

35. A method of targeting an antigen to an epithelial cell in a mammal, comprising linking said antigen to an Ssp or active fragment thereof to produce a Ssp chimeric antigen and administering said chimeric antigen to said mammal.

36. The method of claim 35, wherein said Ssp is SspC or SspD.

37. A method of inducing a cytotoxic T cell immune response in a mammal, comprising linking said antigen to an Ssp or active fragment thereof to produce a Ssp chimeric antigen and contacting an antigen-presenting cell with said chimeric antigen.

38. A vaccine comprising a bacterial cell the virulence of which is attenuated by decreased secretion of a Ssp.

39. The vaccine of claim 38, wherein said bacterial cell is a *Salmonella typhimurium* cell.

40. The vaccine of claim 39, wherein said bacterial cell is a *Salmonella enteriditis* cell.

41. The vaccine of claim 38, wherein said bacterial cell is a *Salmonella typhi* cell.

42. A live *Salmonella* cell in which a gene encoding a heterologous antigen is inserted into a Ssp-encoding gene.

43. A method of vaccinating an animal against a *Salmonella* infection comprising administering the vaccine of claim 38.

44. A substantially pure StpA polypeptide.

45. A method of dephosphorylating a protein, comprising contacting said protein with the polypeptide of claim 44 or an active fragment thereof.

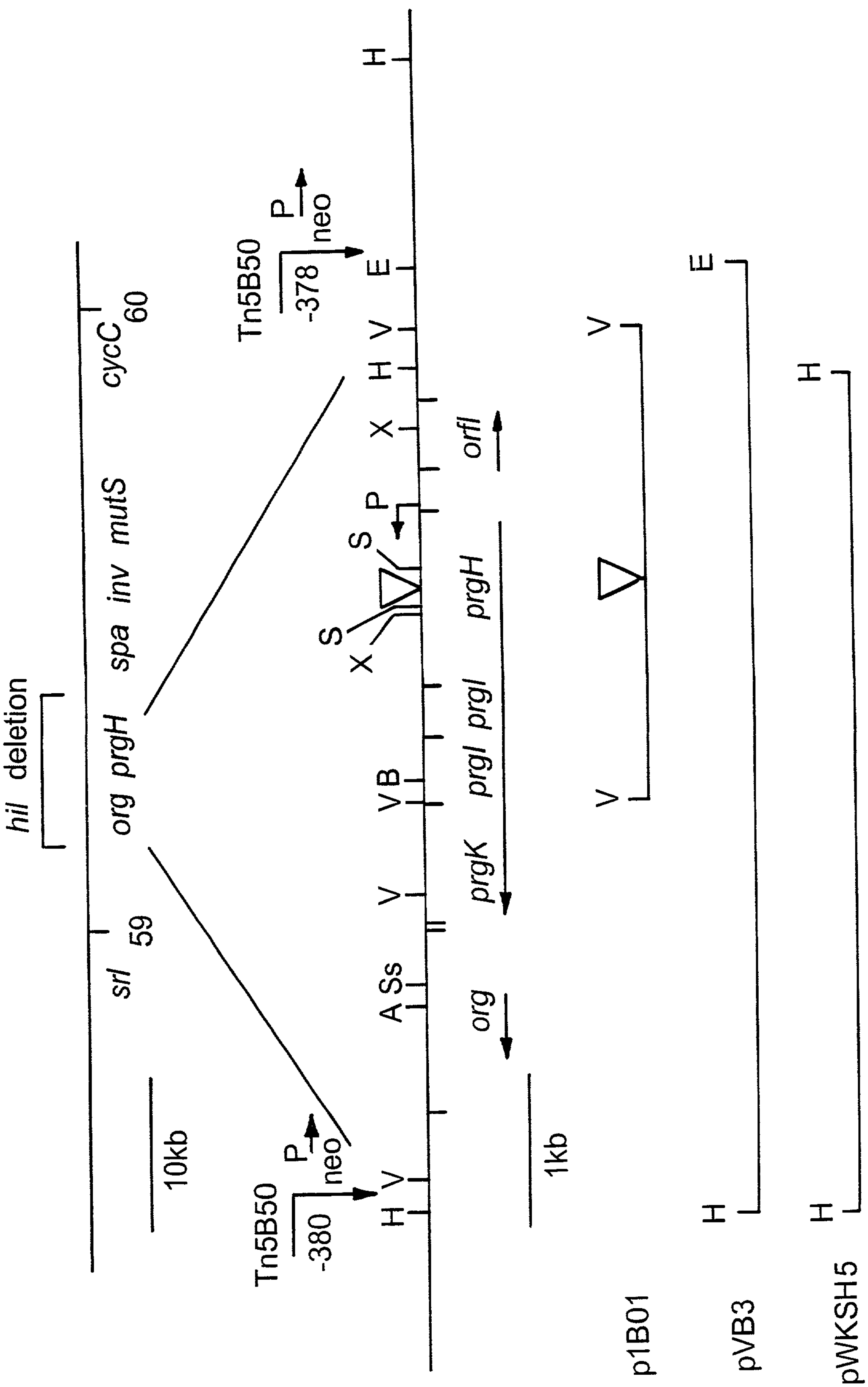

hil deletion
srl 59
org prgH
spa inv mutS
cycC 60
10kb
Tn5B50
-380
-378
P neo
H V
A Ss
V
V B
X S
S
P
X
H V
E
H
org
prgK
prgJ prgI
prgH
orfI
1kb
V
V
E
H
H
H
p1B01
pVB3
pWKSH5

FIG. 1

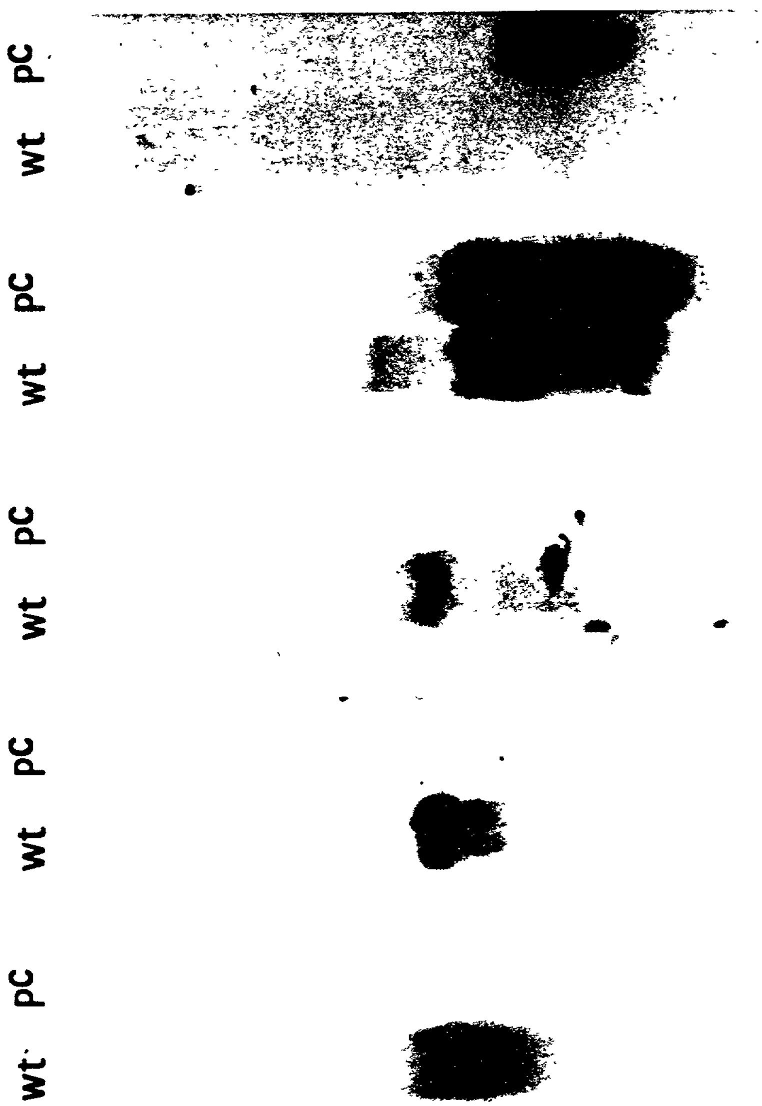


FIG. 2A FIG. 2B FIG. 2C FIG. 2D FIG. 2E

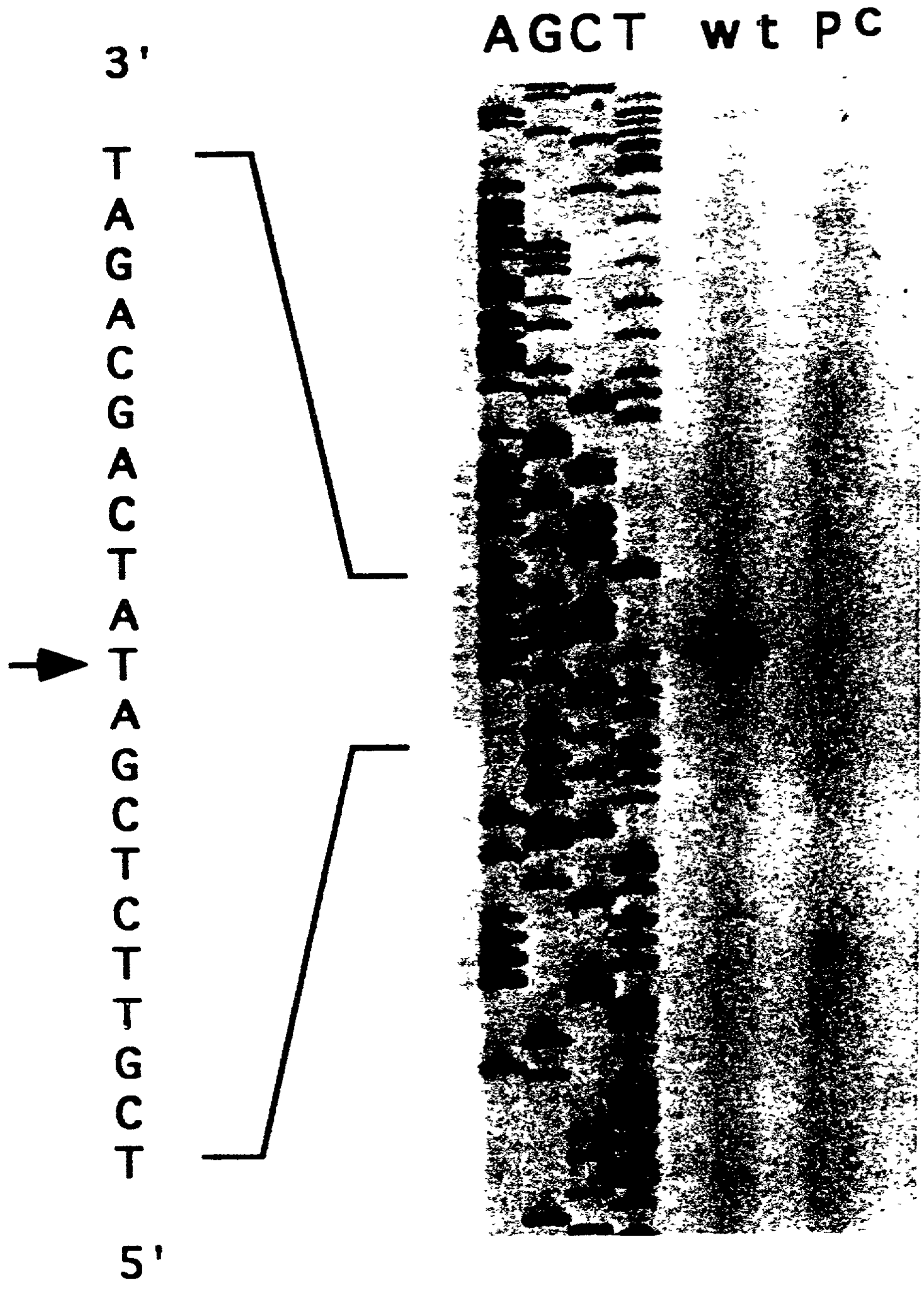


FIG. 3

```
                  *  **             +* **   ** * * **   ** *****+*   +* +
PrgI   1   ..MATPWSGY........LDDVSAKFDTGVDNLQTQVTEALDKLAAKPSDP  41
MxiH   1   MSVTVPNDDWT.......LSSLSETFDDGTQTLQGELTLALDKLAKNPSNP  44
YscF   1   ...MSNFSGFTKGNDIADLDAVAQTLKKPADDANKAVNDSIAALKDTPDNP  48

           *+++ ++******* *****+*****+** ++ ****+*+*+*
PrgI  42   ALLAAYQSKLSEYNLYRNAQSNTVKVFKDIDAAIIQNFR   80 SEQ ID NO:32:
MxiH  45   QLLAEYQSKLSEYTLYRNAQSNTVKVIKDVDAAIIQNFR   83 SEQ ID NO:33:
YscF  49   ALLADLQHSINKWSVIYNISSTIVRSMKDLMQGILQKFP   87 SEQ ID NO:34:
```

FIG. 4A

```
                + + + + +  +    + +  + ++ +        +      +
PrgJ   1   MSIATIVPENAV.IGQAVNIRSMETDIVSLDDRLLQAFSGSAIATAVDKQT  50
MxiI   1   ..MNYIYPVNQVDIIKASDFQSQE..ISSLEDVVSAKYSDIKMDTDIQVSQ  47

           +     +    + ++  +    + +   ++   ++ ++ + ++++++ +    SEQ ID
PrgJ  51   ITNRIEDPNLVTDPKELAISQEMISDYNLYVSMVSTLTRKGVGAVETLLRS 101 NO:35
MxiI  48   IMEMVSNPESL.NPESLAKLQTTLSNYSIGVSLAGTLARKTVSAVETLLKS  97 SEQ ID
                                                                    NO:36
```

FIG. 4B

```
            ***  * *  +***+ ++*  *+  ** + **++ ***+  *** + +
PrgK   1   .MIRRYLYTFLLVMTLAGCKDK.DLLKGLDQEQANEVIAVLQMHNIEANKI  49
MxiJ   1   .MIRYKGFILFLLLMLIGCEQREELISNLSQRQANEIISVLERHNITARKV  50
YscJ   1   MKVKTSLSTLILILFLTGCK..VDLYTGISQKEGNEMLALLRQEGLSADKE  49

           * ** * *  + * **  +*   * + ++   ** * ***+ +*******
PrgK  50   DSGKLGYSITVAEPDFTAAVYWIKTYQLPPRPRVEIAQMFPADSLVSSPR   99
MxiJ  51   DGGKQGISVQVEKGTFASAVDLMRMYDLPNPERVDISQMFPTDSLVSSPR  100
YscJ  50   PDKDGKIKLLVEESDVAQAIDILKRKGYPHESFSTLQDVFPKDGLISSPI   99

           *+*+*+*+*+*+*+*+*+      ++*+*+*+ ***    * *   ** * +
PrgK 100   AEKARLYSAIEQRLEQSLQTMEGVLSARVHISYDIDAGENGRPPKPVHLS  149
MxiJ 101   AEKARLYSAIEQRLEQSLVSIGGVISAKIHVSYDLE..EKNISSKPMHIS  148
YscJ 100   EELARLNYAKAQEISRTLSEIDGVLVARVHVVLPEEQNNKGKKGVAASAS  149

              * *         ** ++*****+** ** +*****+ +       *     *
PrgK 150   ALAVYERGSPLAHQISDIKRFLKNSFADVDYDNISVVL....SERSDAQL  195
MxiJ 149   VIAIYDSPKESELLVSNIKRFLKNTFSDVKYENISVIL....TPKEEYVY  195
YscJ 150   VFIKHAADIQFDTYIPQIKQLVNNSIEGLAYDRISVILVPSVDVRQSSHL  199

              *   **       * * +* ****       +*   *       +
PrgK 196   Q..APGTPVKRNSFATSWIVLIILLSVMSAGFGVWYYKNHYARNKKGITA  243 SEQ ID
MxiJ 196   TNVQPVKEVKSEFLTNEVIYLFLGMAVLVVILLVWAFKTGWFKRNKI    242 NO:37
YscJ 200   P..RNTSILSIQVSEESKGRLIGLLSLLILLLPVTNLAQYFWLQRKK    244 SEQ ID
                                                                   NO:38
PrgK 243   DDKAKSSNE   252  SEQ ID NO:39
```

FIG. 4C

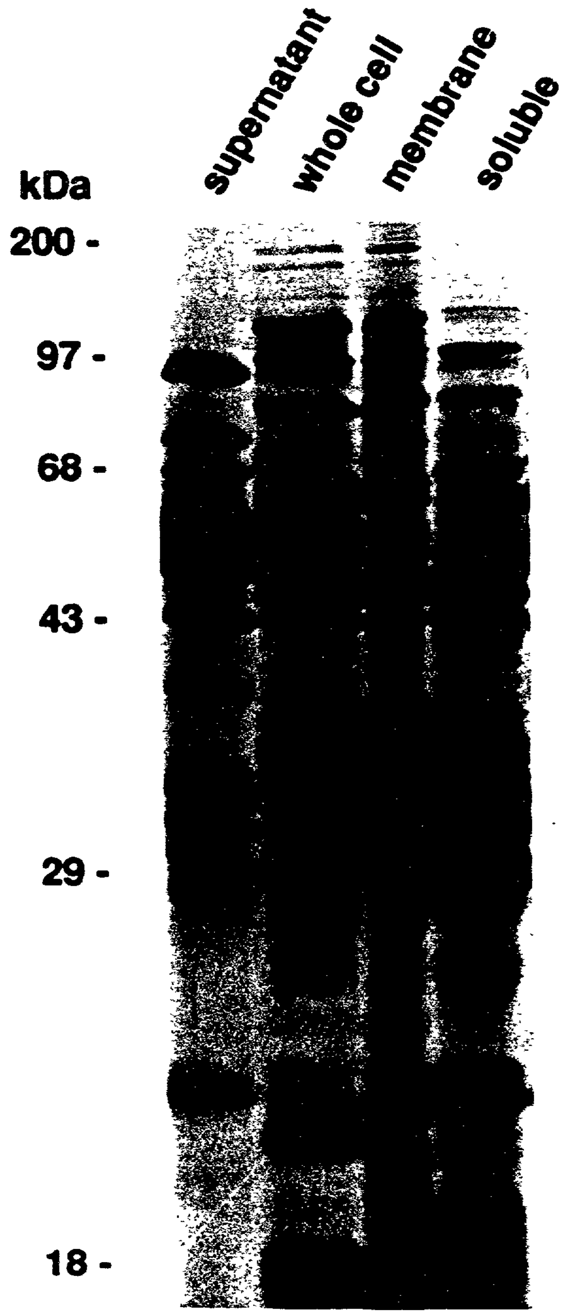

supernatant
whole cell
membrane
soluble
kDa
200 -
97 -
68 -
43 -
29 -
18 -

FIG. 5

6/24

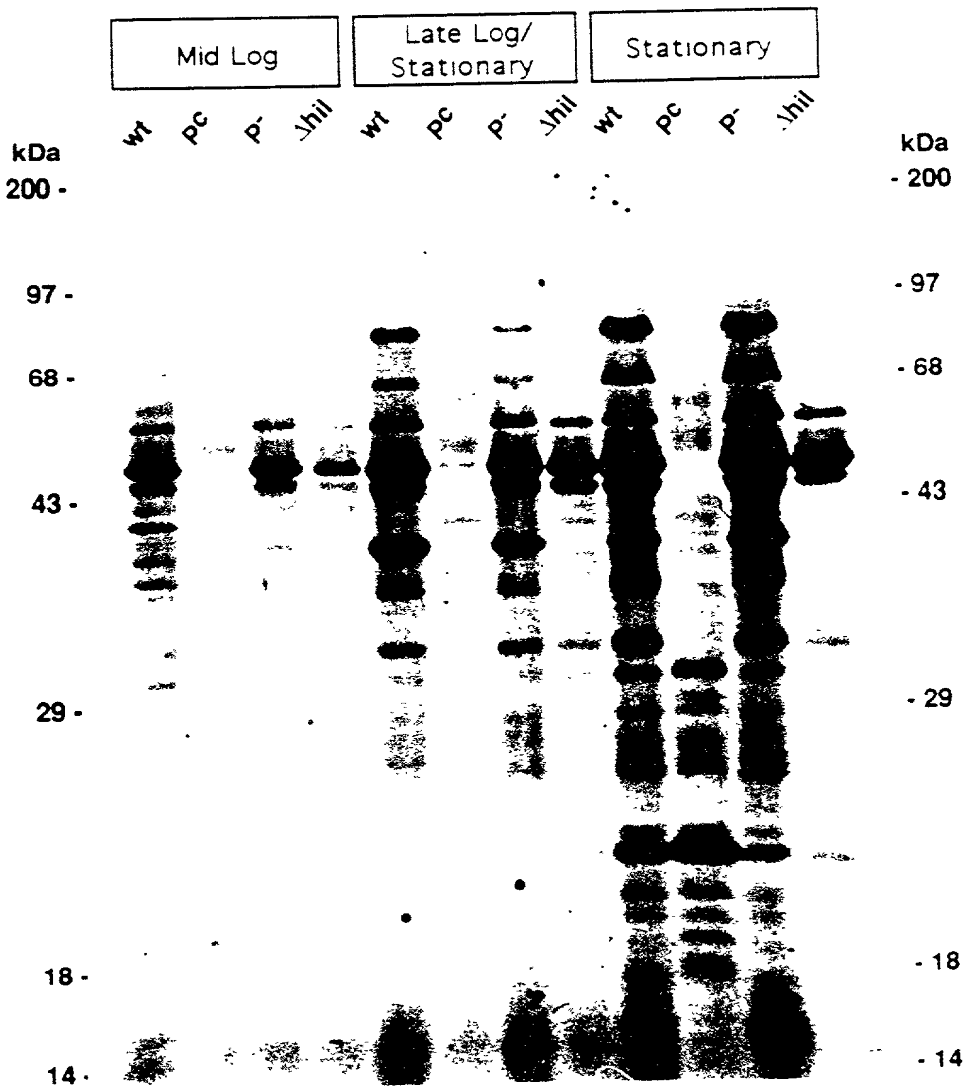

Mid Log
Late Log/ Stationary
Stationary
wt
pc
P-
hil
wt
pc
P-
hil
wt
pc
P-
hil
kDa
200 -
97 -
68 -
43 -
29 -
18 -
14 -
kDa
- 200
- 97
- 68
- 43
- 29
- 18
- 14

FIG. 6

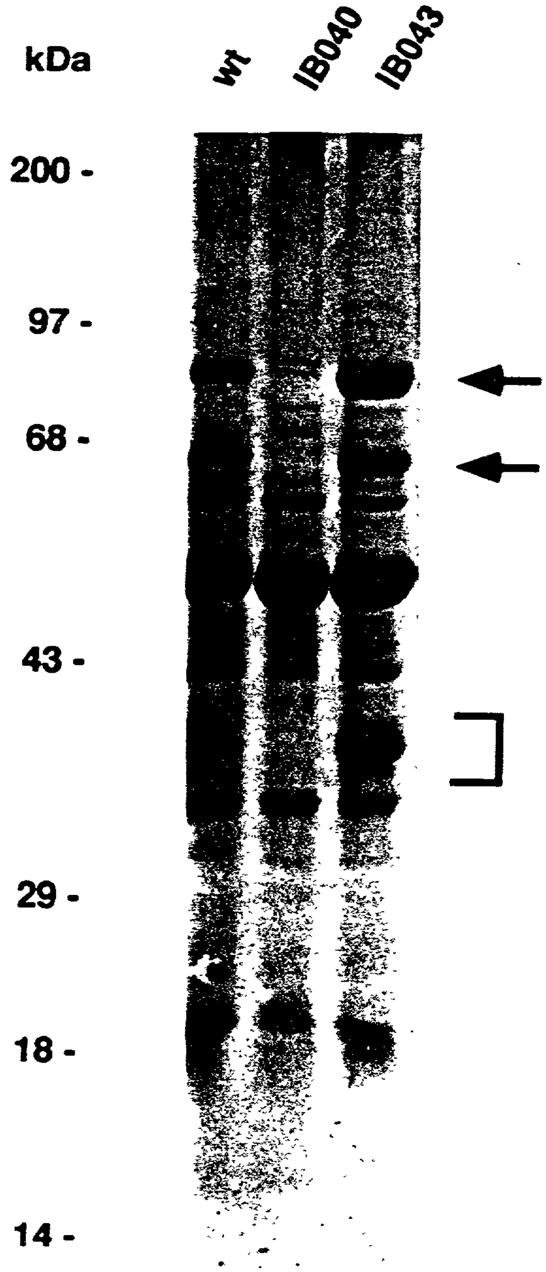


FIG. 7

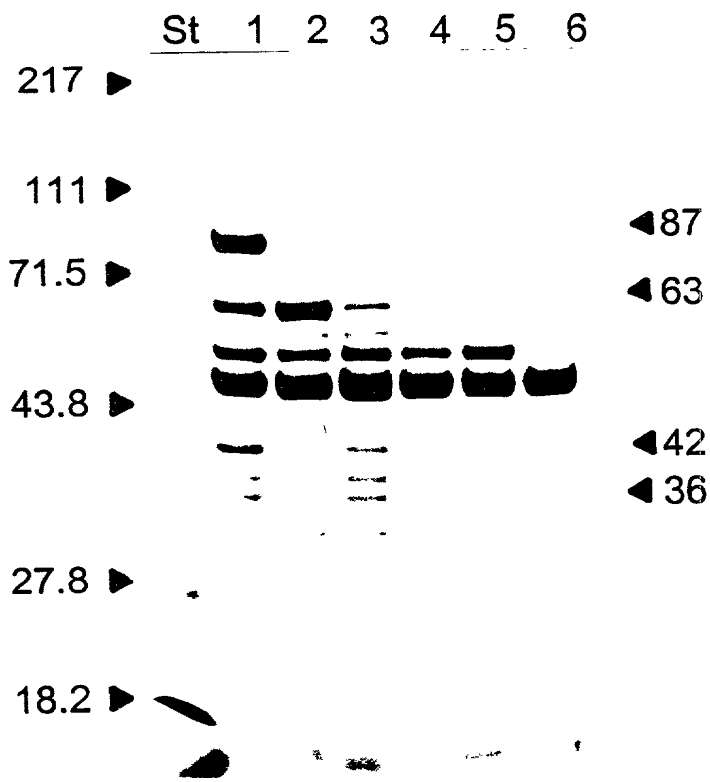
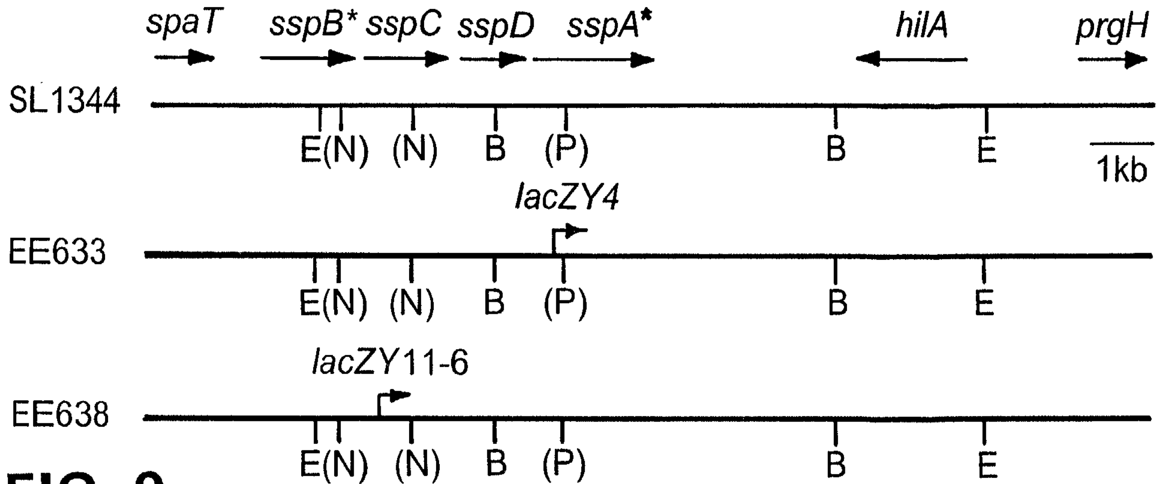
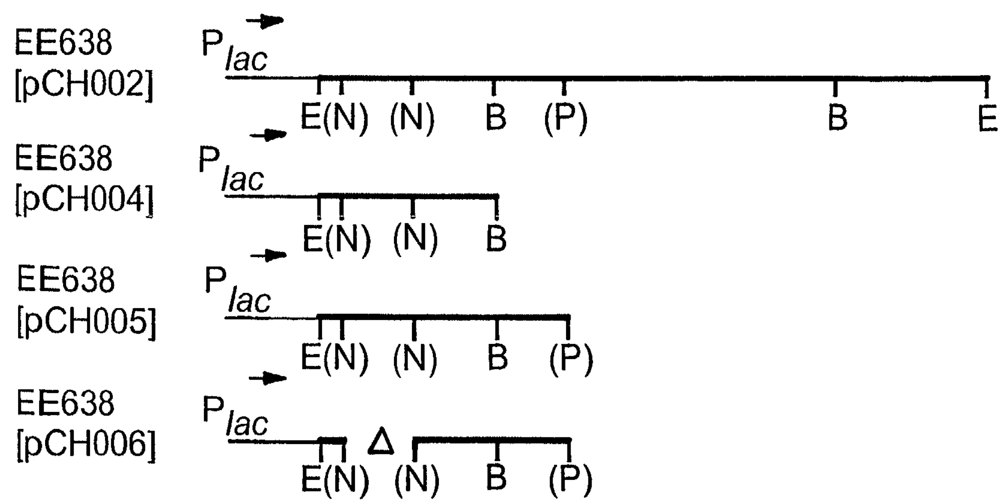


FIG. 8



	INVASION [% of inoculum]	Ssp A (87)	Ssp C (42)	Ssp D (36)
SL1344	17.5 ± 2.6	+	+	+
EE633	23.2 ± 3.5	–	+	+
EE638	0.1 ± 0.05	–	–	–

FIG. 9



	INVASION [% of inoculum]	Ssp A (87)	Ssp C (42)	Ssp D (36)
EE638 [pCH002]	7.1 ± 2.3	+	+	+
EE638 [pCH004]	0.2 ± 0.09	–	+	–
EE638 [pCH005]	16.5 ± 2.7	–	+	+
EE638 [pCH006]	0.1 ± 0.05	-	–	+

FIG. 10

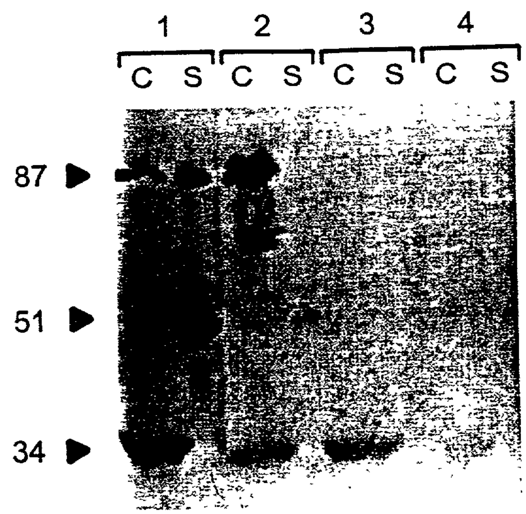


FIG. 11

09/068804

11/24

```
SspB     RKAEETNRIMGCIGKVLGALLTIVSVVAAVFTGGASLALAAVGLAVMVADEIVKAATGVSFIQQALNPIMEHVLKPLMELIGKAITKALEGLGVDKKTAE
IpaB 298 RKAEELNRVMGCVGKIIGALLTIVSVVAAAFSGGASLALAAVGLALMVTDAIVQAATGNSFMEQALNPIMKAVIEPLIKLLSDAFTKMLEGLGVDSKKAK

SspB     MAGSIVGAIVAAIAMVAVIVVVAVVGKGAAAKLGNALSKMMGETIKKLVPNVLKQLAQNGSKLFTQGMQRITSGLGNVGSKMGLQTNALSKELVGNTLNK
IpaB 398 MIGSILGAIAGALVLVAAVVLVATVGKQAAAKLAENIGKIIGKTLTDLIPKFLKNFSSQLDDLITNAVARLNKFLGAAGDEV......ISKQIISTHLNQ

SspB     VALGMEVTNTAAQSAGGVAEGVFIKNASEALADFMLARFAMDQIQQWLKQSVEIFGENQKVTAELQKAMSSAVQQNADASRFILRQSRA      SEQ ID NO:40
IpaB 492 AVLLGESVNSATQAGGSVASAVFQNSASTNLADLTLSKYQVEQLSKYISEAIEKFGQLQEVIADLLASMSNSQANRTDVAKAILQQTTA 580  SEQ ID NO:41
```

FIG. 12

```
SspC   1          MLISNVGINPAAYLNNHSVENSSQTASQSVSAKDILNSIGI.SSSKVSDLGLSPTLSAPAPGVLTQTPGTITSFLKASIQNTDMNQDLNAL
IpaC   1 MLQKQFCNKLLLDTNKENVMEIQN....TKPTQTLYTDISTKQTQSSSETQKSQNYQQIAAHIPLNVGKNPVLTTTLND.DQLLKLSEQVQHDSEIIARL

SspC  91 ANNVTTKANEVVQTQLREQQAEVGKFFDISGMSSSAVALLAAANTIMLTLNQADSKLSGKLSLVSFDAAKTTASSMMREGMNALSGSISQSALQLGITGV
IpaC  96 TDKKMKDLSEMSHTLTPENT......LDISSLSSNAVSLIISVAVLLSALRTAETKIGSQLSLIAFDATKSAAENIVRQGLAALSSSITGAVTQVGITGI

SspC 191 GAKLEYKGLQNERGALKHNAAKIDKLTTESHSIKNVLNGQNSVKLGAEGVDSLKSLNMKKTGTDATKNLNDATLKSNAGTSATESLGIKDSNKQISPEHQ
IpaC 190 GAKKTHSGISDQKGALRKNLATAQSLEKELAGSKLGLNKQIDTNITSPQTNS...........STKFLGKNKLAPD.............NISLSTEHK

SspC 291 AILSKRLESVESDIRLEQNTMDMTRIDARKMQMTGDLIMKNSVTVGGIAGASGQYAATQERSEQQISQVNNRVASTASDEARESSRKSTSLIQEMLKTME
IpaC 264 TSLSSPDISLQDKIDTQRRTYELNTLSAQQKQNIGRATMETSAVAGNISTSGGRYASALEEEEQLISQASSKQAEEASQVSKEASQATNQLIQKLLNIID

SspC 391 SINQSKASALAAIAGNIRA 409   SEQ ID NO:42
IpaC 364 SINQSKNSAASQIAGNIRA 382   SEQ ID NO:43
```

FIG. 13

```
SspD   1 MLNIQNYSASPHPGIVAERPQTPSASEHVETAVVPSTTEHRGTDIISLSQAATKIHQAQQTLQ...STPPISEENNDERTLARQQLTSSLNALAKSGVSL
IpaD   1  MNITTLTNSISTSSFSPNNTNGSSTETVNSDIKTTTSSHPVSSLTMLNDTLHNIRTTNQALKKELSQKTLTKTSLEEIALHSSQISMDVNKSAQLLDIL

SspD  98 SAEQ...NENLRSAF.SAPTSALFSASPMAQPRTTISDAEIWDMVSQNISAIGDSYLGVYENVVAVYTDFYQAFSDILSKMGGWLLP.GKDGNTVKLDVT
IpaD 100 SRNEYPINKDARELLHSAPKEAELDGDQM......ISHRELWAKIANSINDINEQYLKVYEHAVSSYTQMYQDFSAVLSSLAGWISPGGNDGNSVKLQVN

SspD 193 SLKNDLNSLVNKYNQINSNTVLFPAQSGSGVKVATEAEARQWLSELNLPNSCLKSYGSGYVVTVDLTPLQKMVQDIDGLGAPGKDSKLEMDNAKYQAWQS
IpaD 194 SLKKALEELKEKYK....DKPLYPANNT.....VSQEQANKWLTELGGTIGKVSQKNGGYVVSINMTPIDNMLKSLDNLGGNG...EVVLDNAKYQAWNA

SspD 293 GFKAQEENMKTTLQTLTQKYSNANSLYDNLVKVLSSTISSSLETAKSFLQG 343  SEQ ID NO:44
IpaD 282 GFSAEDETMKNNLQTLVQKYSNANSIFDNLVKVLSSTISSCTDTDKLFLHF 332  SEQ ID NO:45
```

FIG. 14

```
SspA   1 MVTSVRTQPPVIMPGMQTEIKTQATNLAANLSAVRESATATLSGEIKGPQLEDFPALIKQASLD 64   SEQ ID NO:46
```

FIG. 15

09/068804

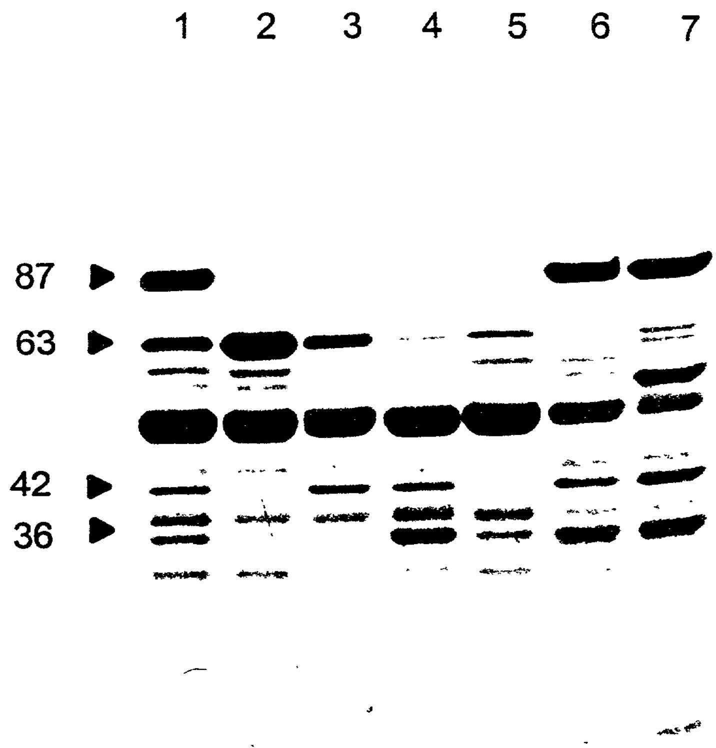


FIG. 16

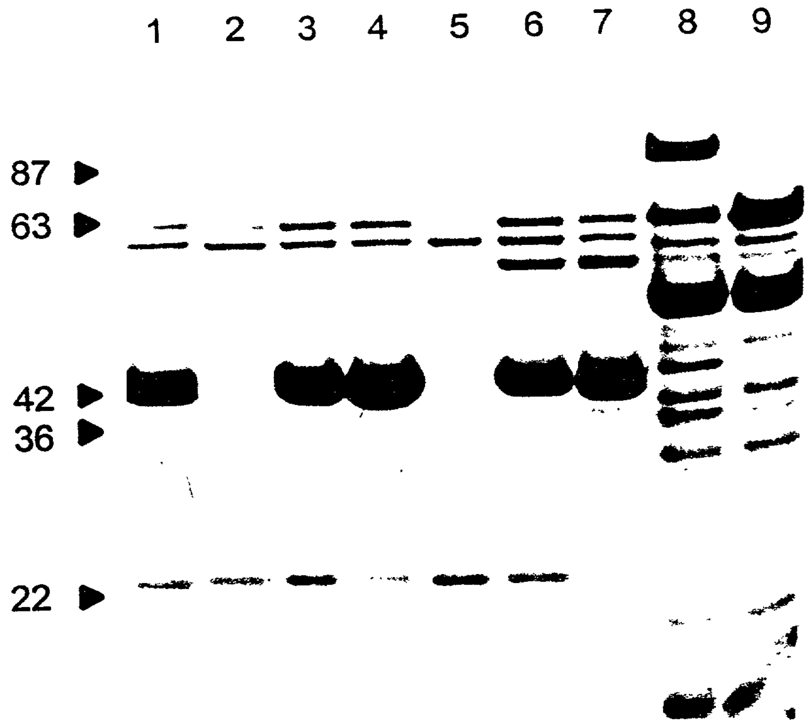


FIG. 17

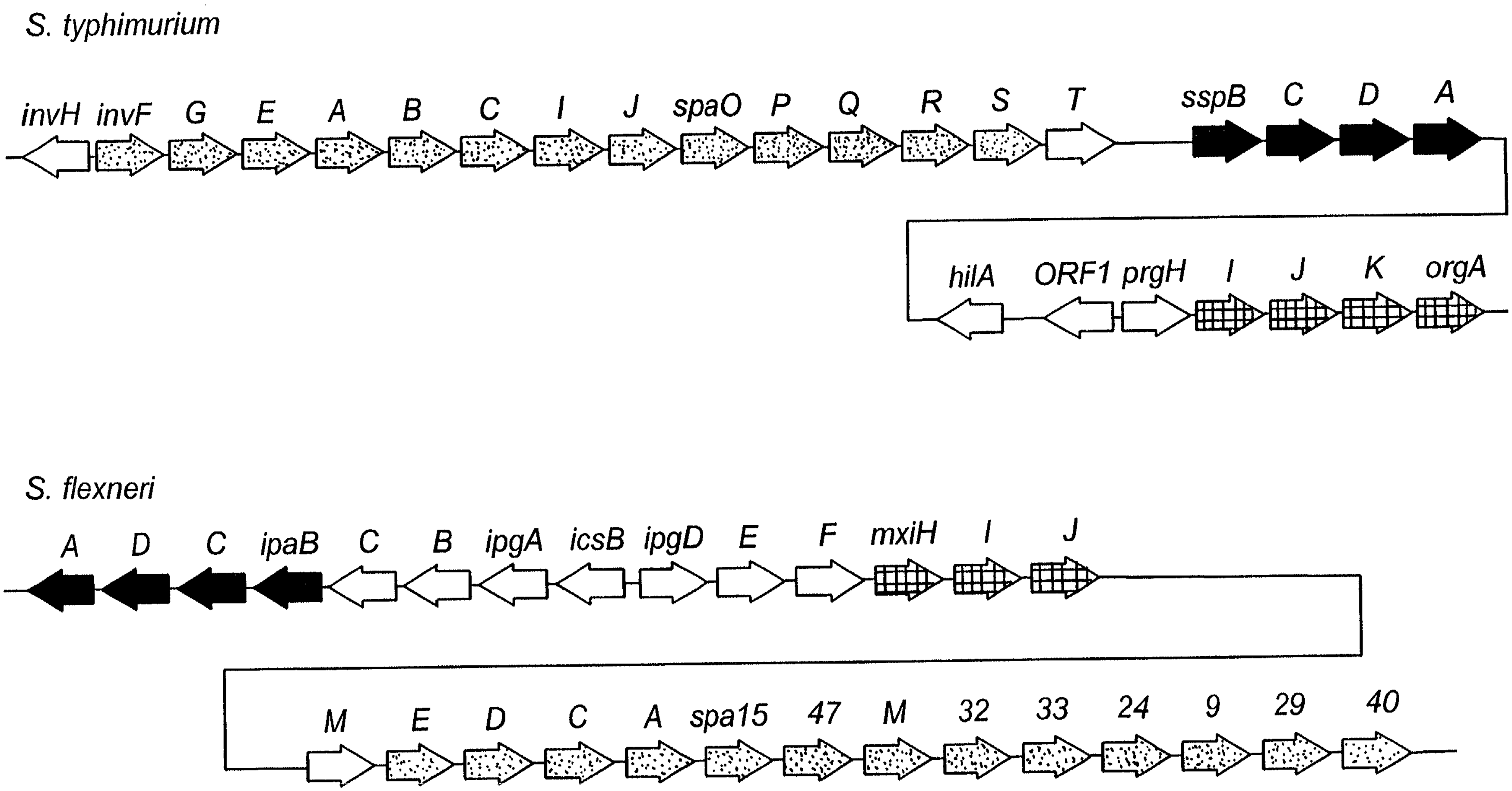

S. typhimurium
invH
invF
G
E
A
B
C
I
J
spaO
P
Q
R
S
T
sspB
C
D
A
hilA
ORF1
prgH
I
J
K
orgA
S. flexneri
A
D
C
ipaB
C
B
ipgA
icsB
ipgD
E
F
mxiH
I
J
M
E
D
C
A
spa15
47
M
32
33
24
9
29
40

FIG. 18

SEQ ID NO:1:

```
CGCAAAGCCG AGGAAACGAA CCGCATTATG GGATGTATCG GGAAAGTCCT CGGCGCGCTG   60
CTAACCATTG TCAGCGTTGT GGCCGCTGTT TTTACCGGTG GGGCGAGTCT GGCGCTGGCT  120
GCGGTGGGAC TTGCGGTAAT GGTGGCCGAT GAAATTGTGA AGGCGGCGAC GGGAGTGTCG  180
TTTATTCAGC AGGCGCTAAA CCCGATTATG GAGCATGTGC TGAAGCCGTT AATGGAGCTG  240
ATTGGCAAGG CGATTACCAA AGCGCTGGAA GGATTAGGCG TCGATAAGAA AACGGCAGAG  300
ATGGCCGGCA GCATTGTTGG TGCGATTGTC GCCGCTATTG CCATGGTGGC GGTCATTGTG  360
GTGGTCGCAG TTGTCGGGAA AGGCGCGGCG GCGAAACTGG GTAACGCGCT GAGCAAAATG  420
ATGGGCGAAA CGATTAAGAA GTTGGTGCCT AACGTGCTGA AACAGTTGGC GCAAAACGGC  480
AGCAAACTCT TTACCCAGGG GATGCAACGT ATTACTAGCG GTCTGGGTAA TGTGGGTAGC  540
AAGATGGGCC TGCAAACGAA TGCCTTAAGT AAAGAGCTGG TAGGTAATAC CCTAAATAAA  600
GTGGCGTTGG GCATGGAAGT CACGAATACC GCAGCCCAGT CAGCCGGTGG TGTTGCCGAG  660
GGCGTATTTA TTAAAAATGC CAGCGAGGCG CTTGCTGATT TTATGCTCGC CCGTTTTGCC  720
ATGGATCAGA TTCAGCAGTG GCTTAAACAA TCCGTAGAAA TATTTGGTGA AAACCAGAAG  780
GTAACGGCGG AACTGCAAAA AGCCATGTCT TCTGCGGTAC AGCAAAATGC GGATGCTTCG  840
CGTTTTATTC TGCGCCAGAG TCGCGCATAA                                    870
```

FIG. 19

SEQ ID NO:2:

```
ATGTTAATTA GTAATGTGGG AATAAATCCC GCCGCTTATT TAAATAATCA TTCTGTTGAG    60
AATAGTTCAC AGACAGCTTC GCAATCCGTT AGCGCTAAAG ATATTCTGAA TAGTATTGGT   120
ATTAGCAGCA GTAAAGTCAG TGACCTGGGG TTGAGTCCTA CACTGAGCGC GCCTGCGCCA   180
GGGGTATTAA CGCAAACCCC CGGAACGATC ACGTCCTTTT TAAAAGCCAG TATTCAAAAT   240
ACCGACATGA ATCAGGATTT GAATGCTCTG GCAAATAATG TCACGACTAA AGCGAATGAG   300
GTTGTGCAAA CCCAGTTACG CGAGCAGCAG GCAGAAGTCG GAAAGTTTTT TGATATTAGC   360
GGAATGTCTT CCAGTGCCGT TGCGCTGTTG GCTGCCGCGA ATACGTTAAT GCTGACGTTG   420
AACCAGGCTG ATAGCAAACT GTCTGGTAAG TTGTCATTAG TCAGTTTTGA TGCAGCTAAA   480
ACGACGGCAA GCTCCATGAT GCGCGAAGGG ATGAATGCGT TGTCCGGTAG TATTTCCCAG   540
AGCGCGCTTC AGTTGGGGAT CACTGGCGTG GGCGCCAAAC TGGAATATAA GGGGCTGCAG   600
AATGAAAGAG GCGCGCTTAA ACATAATGCC GCGAAGATCG ATAAACTGAC CACTGAAAGC   660
CACAGTATTA AAAACGTGCT GAACGGGCAG AATAGCGTCA AACTCGGTGC TGAAGGCGTC   720
GATTCTCTGA AATCGTTAAA TATGAAGAAA ACCGGTACCG ATGCGACGAA AAATCTTAAT   780
GATGCGACGC TTAAATCTAA TGCCGGAACC AGCGCCACGG AAAGTCTGGG TATTAAAGAC   840
AGTAATAAAC AAATCTCCCC TGAACATCAG GCTATTCTGT CGAAACGTCT TGAGTCTGTC   900
GAATCCGATA TTCGTCTTGA GCAGAATACC ATGGATATGA CCCGAATCGA TGCGCGCAAG   960
ATGCAGATGA CGGGCGATCT GATTATGAAG AACTCGGTCA CGGTCGGTGG TATTGCAGGG  1020
GCGTCCGGGC AGTACGCCGC TACTCAGGAA CGTTCCGAGC AGCAAATTAG CCAGGTGAAT  1080
AACCGGGTTG CCAGCACCGC ATCGGACGAA GCCCGTGAAA GTTCACGTAA ATCGACCAGC  1140
CTGATTCAGG AAATGCTGAA AACAATGGAG AGCATTAACC AGTCGAAAGC ATCCGCACTC  1200
GCTGCTATCG CAGGCAATAT TCGCGCTTAA                                   1230
```

FIG. 20

SEQ ID NO:3:

ATGCTTAATA TTCAAAATTA TTCCGCTTCT CCTCATCCGG GGATCGTTGC CGAACGGCCG 60
CAGACTCCCT CGGCGAGCGA GCACGTCGAG ACTGCCGTGG TACCGTCTAC CACAGAACAT 120
CGCGGTACAG ATATCATTTC ATTATCGCAG GCGGCTACTA AAATCCACCA GGCACAGCAG 180
ACGCTGCAGT CAACGCCACC GATCTCTGAA GAGAATAATG ACGAGCGCAC GCTGGCGCGC 240
CAGCAGTTGA CCAGCAGCCT GAATGCGCTG GCGAAGTCCG GCGTGTCATT ATCCGCAGAA 300
CAAAATGAGA ACCTGCGGAG CGCGTTTTCT GCGCCGACGT CGGCCTTATT TAGCGCTTCG 360
CCTATGGCGC AGCCGAGAAC AACCATTTCT GATGCTGAGA TTTGGGATAT GGTTTCCCAA 420
AATATATCGG CGATAGGTGA CAGCTATCTG GGCGTTTATG AAAACGTTGT CGCAGTCTAT 480
ACCGATTTTT ATCAGGCCTT CAGTGATATT CTTTCCAAAA TGGGAGGCTG GTTATTACCA 540
GGTAAGGACG GTAATACCGT TAAGCTAGAT GTTACCTCAC TCAAAAATGA TTTAAACAGT 600
TTAGTCAATA AATATAATCA AATAAACAGT AATACCGTTT TATTTCCAGC GCAGTCAGGC 660
AGCGGCGTTA AAGTAGCCAC TGAAGCGGAA GCGAGACAGT GGCTCAGTGA ATTGAATTTA 720
CCGAATAGCT GCCTGAAATC TTATGGATCC GGTTATGTCG TCACCGTTGA TCTGACGCCA 780
TTACAAAAAA TGGTTCAGGA TATTGATGGT TTAGGCGCGC CGGGAAAAGA CTCAAAACTC 840
GAAATGGATA ACGCCAAATA TCAAGCCTGG CAGTCGGGTT TTAAAGCGCA GGAAGAAAAT 900
ATGAAAACCA CATTACAGAC GCTGACGCAA AAATATAGCA ATGCCAATTC ATTGTACGAC 960
AACCTGGTAA AAGTGCTGAG CAGTACGATA AGTAGCAGCC TGGAAACCGC CAAAAGCTTC 1020
CTGCAAGGAT AA 1032

FIG. 21

ATGGTTACAA GTGTAAGGAC TCAGCCCCCC GTCATAATGC CAGGTATGCA GACCGAGATC
AAAACGCAGG CCACGAATCT TGCGGCGAAT CTTTCCGCAG TCAGAGAAAG TGCCACAGCG
ACGCTGTCAG GGGAAATTAA AGGCCCGCAA CTGGAAGATT TTCCCGCGCT GATCAAACAG
GCGAGTCTGG ATGC SEQ ID NO:4

RKAEETNRIMGCIGKVLGALLTIVSVVAAVFTGGASLALAAVGLAVMVADEIVKAATGVS
FIQQALNPIMEHVLKPLMELIGKAITKALEGLGVDRKRQRWPAALLVRLSPLCHGDAVIV
VVAVVGKGAAAKLGNALSKMMGETIKKLVPNVLKQLAQNGSKLFTQGMQRITSGLGNVGS
KMGLQTNALSKELVGNTLNKVALGMEVTNTAAQSAGGVAEGVFIKNASEALADFMLARFA
MDQIQQWLKQSVEIFGENQKVTAELQKAMSSAVQQNADASRFILRQSRAZ SEQ ID NO:5

FIG. 22

MLISNVGINPAAYLNNHSVENSSQTASQSVSAKDILNSIGISSSKVSDLGLSPTLSAPAP
GVLTQTPGTITSFLKASIQNTDMNQDLNALANNVTTKANEVVQTQLREQQAEVGKFFDIS
GMSSSAVALLAAANTLMLTLNQADSKLSGKLSLVSFDAAKTTASSMMREGMNALSGSISQ
SALQLGITGVGAKLEYKGLQNERGALKHNAAKIDKLTTESHSIKNVLNGQNSVKLGAEGV
DSLKSLNMKKTGTDATKNLNDATLKSNAGTSATESLGIKDSNKQISPEHQAILSKRLESV
ESDIRLEQNTMDMTRIDARKMQMTGDLIMKNSVTVGGIAGASGQYAATQERSEQQISQVN
NRVASTASDEARESSRKSTSLIQEMLKTMESINQSKASAL AAIAGNIRAZ
SEQ ID NO:6

MLNIQNYSASPHPGIVAERPQTPSASEHVETAVVPSTTEHRGTDIISLSQAATKIHQAQQ
TLQSTPPISEENNDERTLARQQLTSSLNALAKSGVSLSAEQNENLRSAFSAPTSALFSAS
PMAQPRTTISDAEIWDMVSQNISAIGDSYLGVYENVVAVYTDFYQAFSDILSKMGGWLLP
GKDGNTVKLDVTSLKNDLNSLVNKYNQINSNTVLFPAQSGSGVKVATEAEARQWLSELNL
PNSCLKSYGSGYVVTVDLTPLQKMVQDIDGLGAPGKDSKLEMDNAKYQAWQSGFKAQEEN
MKTTLQTLTQKYSNANSLYDNLVKVLSSTISSSLETAKSFLQGZ
SEQ ID NO:7

MVTSVRTQPPVIMPGMQTEIKTQATNLAANLSAVRESATATLSGEIKGPQLEDFPALIKQ
ASLD
SEQ ID NO:8

FIG. 23

ATGCATTATTTTTTTATCATCGTAATCTGGTTGCTTAGCATAAATACGGCATGGGCTGAT
TCTGGCTTCAGGCTGAAAAAATGTTCAATATTGAATCCGAACTACTTTACGCTATCGCCCAGCAG
GAATCGGCGATGAAACCTGGCGCCATTGGTCATAACCGAGATGGTTCAACCGATCTTGGCCTGAT
GCAAATTAACAGCTTCCATATGAAAAGGCTGAAAAAAATGGGGATTAGTGAAAAACAGTTGTTAC
AGGACCCCTGCATTTCTGTCATTGTGGGCGACCTCCATTTTATCAGATATGATGAAAATCTACGG
TTATAGCTGGGAGGCCGTTGGCGCTTATAATGCCGGGACGTCGCCGAAACGATCGGATATAAGGA
AACGTTATGCTAAAAAAATTTGGGAGAATTACAGAAAATTAAAAGGAATGTCAGCAGAAGAGAAA
AACAAAAGACTTTCTATCGCGGCAAACAAATAA (SEQ ID NO:9)

ATCAGCTTGCCGTCGTCATAAGCAACTGGGCTTGCATTGCTTTTAGTTGTACAAACTGTGCAGGC
GTCTTCCAGCATTCTATTGTTCCGTGAATCCGGAAATCTGCACGTACCTGCTCCAGATTACTATG
AGGATTATCCTTAAGTACAAGGGCCGCCGCCATCGTTCCGGTTCTTCCCACTCCGCCCAGACAAT
GAATCATCGGTAAATGCTTATCTGATGAACTACGCCCCGGCGCGCCATTTTGGTTACTATTTTTC
ACCCTATCCGCCAGGTATTCTAACTGATCCGTAGACGGTAACGGCTGGTGATCTGGCCAATTTTT
CACATGCAATACCGGGATTGTATACCGCTTTCCCCGCAGGACAGTTGCATATTGTATTGGTCTAT
CGCTTCTCCCTGACTGGCTGAGCTCTCTTTTTGGCTGTTGGTATGCACCTCGCCAAAGGTGTAGC
TCCCTCTGAAATAGGTGGTAATTGTTTTGCCTGCATCTGATCTTCCGACGTTAACACCACCAGGC
ACGAGCATTCTTTTTCAAGAAGCATTTTCATATGCGCTTCCAGCGCATCCCGGCGATTT
(SEQ ID NO:10)

FIG. 24

MHYFFIIVIWLLSINTAWADSGFRLKKCSILNPNYFTLSPSRNRRZNLAPLVITEMVQPI
LAZCKLTASIZKGZKKWGLVKNSCYRTPAFLSLWATSILSDMMKIYGYSWEAVGAYNAGT
SPKRSDIRKRYAKKIWENYRKLKGMSAEEKNKRLSIAANK

SEQ ID NO:11

WPGTICGQQHSINQQTQVKLSDGMPVPVIRLTFDGKPVALAGIRTQKIRPDRLEAHMKML
LEKECSCLVVLTSERSDAGKTITTYFRGSYTFGEVHTNSQKVSSASQGEAIDQYNMQLSC
GEKRYTIPVLHVKNWPDHQPLPSTDQLEYLADRVKNSNQNGAPGRSSSDKHLPMIHCLGG
VGRTGTMAAALVLKDNPHSNLEQVRADFRIHGTIECWKTPAQFVQLKAMQAQLLMTTAS

SEQ ID NO:12

FIG. 25

MRDCLNNGNPVLNVGASGLTTLPDRLPPHITTLVIPDNNLTSLPELPEGLRELEVSGNLQ
LTSLPSLPQGLQKLWAYNNWLASLPTLPPGLGDLAVSNNQLTSLPEMPPALRELRVSGNN
LTSLRALPSGLQKLWAYNNRLTSLPEMSPGLQELDVSHNQLTRLPQSLTGLSSAARVYLD
GNPLSVRTRDRLCGHHWPFRHQDTLRYGGAFRPREARALHLAVADWLTSAREGEAAQADR
WQAFGLEDNAAAFSLVLDRLRETENFKKDAGFKAQISSWLTQLAEDAALRAKTFAMATEA
TSTCEDRVTHALHQMNNVQLVHNAEKGEYDNNLQGLVSTGREMFRLATLEQIAREKAGTL
ALVDDVEVYLAFQNKLKESLELTSVTSEMRFFDVSGVTVSDLQAADVQVKTAENSGFSKW
ILQWGPLHSVLERKVPERFNALREKQISDYEDTYRKLYDEVLKSSGLVDDTDAERTIGVS
AMDSAKKEFLDGLRALVDEVLGSYLTARWRLNZ

SEQ ID NO:14

FIG. 27

ATGCGTGATT GCCTGAATAA CGGCAATCCA GTGCTTAACG TGGGAGCGTC AGGTCTTACC
ACCTTACCAG ACCGTTTACC ACCGCATATT ACAACACTGG TTATTCCTGA TAATAATCTG
ACCAGCCTGC CGGAGTTGCC GGAAGGACTA CGGGAGCTGG AGGTCTCTGG TAACCTACAA
CTGACCAGCC TGCCATCGCT GCCGCAGGGA CTACAGAAGC TGTGGGCCTA TAATAATTGG
CTGGCCAGCC TGCCGACGTT GCCGCCAGGA CTAGGGGATC TGGCGGTCTC TAATAACCAG
CTGACCAGCC TGCCGGAGAT GCCGCCAGCA CTACGGGAGC TGAGGGTCTC TGGTAACAAC
CTGACCAGCT GCGCGCGCTG CCGTCAGGAC TACAGAAGCT GTGGGCCTAT AATAATCGGC
TGACCAGCCT GCCGGAGATG TCGCCAGGAC TACAGGAGCT GGATGTCTCT CATAACCAGC
TGACCCGCCT GCCGCAAAGC CTCACGGGTC TGTCTTCAGC GGCACGCGTA TATCTGGACG
GGAATCCACT GTCTGTACGC ACTCGTGACA GGCTCTGCGG ACATCATTGG CCATTCAGGC
ATCAGGATAC ACTTCGATAT GGCGGGGCCT TCCGTCCCCG GGAAGCCCGG GCACTGCACC
TGGCGGTCGC TGACTGGCTG ACGTCTGCAC GGGAGGGGGA AGCGGCCCAG GCAGACAGAT
GGCAGGCGTT CGGACTGGAA GATAACGCCG CCGCCTTCAG CCTGGTCCTG GACAGACTGC
GTGAGACGGA AAACTTCAAA AAAGACGCGG GCTTTAAGGC ACAGATATCA TCCTGGCTGA
CACAACTGGC TGAAGATGCT GCGCTGAGAG CAAAAACCTT TGCCATGGCA ACAGAGGCAA
CATCAACCTG CGAGGACCGG GTCACACATG CCCTGCACCA GATGAATAAC GTACAACTGG
TACATAATGC AGAAAAAGGG GAATACGACA ACAATCTCCA GGGGCTGGTT TCCACGGGGC
GTGAGATGTT CCGCCTGGCA ACACTGGAAC AGATTGCCCG GGAAAAAGCC GGAACACTGG
CTTTAGTCGA TGACGTTGAG GTCTATCTGG CGTTCCAGAA TAAGCTGAAG GAATCACTTG
AGCTGACCAG CGTGACGTCA GAAATGCGTT TCTTTGACGT TTCCGGCGTG ACGGTTTCAG
ACCTTCAGGC TGCGGACGTT CAGGTGAAAA CCGCTGAAAA CAGCGGGTTC AGTAAATGGA
TACTGCAGTG GGGGCCGTTA CACAGCGTGC TGGAACGCAA AGTGCCGGAA CGCTTTAACG
CGCTTCGTGA AAAGCAAATA TCGGATTATG AAGACACGTA CCGGAAGCTG TATGACGAAG
TGCTGAAATC GTCCGGGCTG GTCGACGATA CCGATGCAGA ACGTACTATC GGAGTAAGTG
CGATGGATAG TGCGAAAAAA GAATTTCTGG ATGGCCTGCG CGCTCTTGTG GATGAGGTGC
TGGGTAGCTA TCTGACAGCC CGGTGGCGTC TTAACTGA

SEQ ID NO:13

FIG. 26

CGCAAAGCCGAGGAAACGAACCGCATTATGGGATGTATCGGGAAAGTCCTCGGCGCGCTG
CTAACCATTGTCAGCGTTGTGGCCGCTGTTTTTACCGGTGGGCGAGTCTGGCGCTGGCT
GCGGTGGGACTTGCGGTAATGGTGGCCGATGAAATTGTGAAGGCGGCGACGGGAGTGTCG
TTTATTCAGCAGGCGCTAAACCCGATTATGGAGCATGTGCTGAAGCCGTTAATGGAGCTG
ATTGGCAAGGCGATTACCAAAGCGCTGGAAGGATTAGGCGTCGATAAGAAAACGGCAGAG
ATGGCCGGCAGCATTGTTGGTGCGATTGTCGCCGCTATTGCCATGGTGGCGGTCATTGTG
GTGGTCGCAGTTGTCGGGAAAGGCGCGGCGGCGAAACTGGGTAACGCGCTGAGCAAAATG
ATGGGCGAAACGATTAAGAAGTTGGTGCCTAACGTGCTGAAACAGTTGGCGCAAAACGGC
AGCAAACTCTTTACCCAGGGGATGCAACGTATTACTAGCGGTCTGGGTAATGTGGGTAGC
AAGATGGGCCTGCAAACGAATGCCTTAAGTAAAGAGCTGGTAGGTAATACCCTAAATAAA
GTGGCGTTGGGCATGGAAGTCACGAATACCGCAGCCCAGTCAGCCGGTGGTGTTGCCGAG
GGCGTATTTATTAAAAATGCCAGCGAGGCGCTTGCTGATTTTATGCTCGCCCGTTTTGCC
ATGGATCAGATTCAGCAGTGGCTTAAACAATCCGTAGAAATATTTGGTGAAAACCAGAAG
GTAACGGCGGAACTGCAAAAAGCCATGTCTTCTGCGGTACAGCAAAATGCGGATGCTTCG
CGTTTTATTCTGCGCCAGAGTCGCGCATAAAAACTGCCAAAATAAAGGGAGAAAAAT**ATG**
TTAATTAGTAATGTGGGAATAAATCCCGCCGCTTATTTAAATAATCATTCTGTTGAGAAT
AGTTCACAGACAGCTTCGCAATCCGTTAGCGCTAAAGATATTCTGAATAGTATTGGTATT
AGCAGCAGTAAAGTCAGTGACCTGGGGTTGAGTCCTACACTGAGCGCGCCTGCGCCAGGG
GTATTAACGCAAACCCCCGGAACGATCACGTCCTTTTTAAAAGCCAGTATTCAAAATACC
GACATGAATCAGGATTTGAATGCTCTGGCAAATAATGTCACGACTAAAGCGAATGAGGTT
GTGCAAACCCAGTTACGCGAGCAGCAGGCAGAAGTCGGAAAGTTTTTTGATATTAGCGGA
ATGTCTTCCAGTGCCGTTGCGCTGTTGGCTGCCGCGAATACGTTAATGCTGACGTTGAAC
CAGGCTGATAGCAAACTGTCTGGTAAGTTGTCATTAGTCAGTTTTGATGCAGCTAAAACG
ACGGCAAGCTCCATGATGCGCGAAGGGATGAATGCGTTGTCCGGTAGTATTTCCCAGAGC
GCGCTTCAGTTGGGGATCACTGGCGTGGGCGCCAAACTGGAATATAAGGGGCTGCAGAAT
GAAAGAGGCGCGCTTAAACATAATGCCGCGAAGATCGATAAACTGACCACTGAAAGCCAC
AGTATTAAAAACGTGCTGAACGGGCAGAATAGCGTCAAACTCGGTGCTGAAGGCGTCGAT
TCTCTGAAATCGTTAAATATGAAGAAAACCGGTACCGATGCGACGAAAAATCTTAATGAT
GCGACGCTTAAATCTAATGCCGGAACCAGCGCCACGGAAAGTCTGGGTATTAAAGACAGT
AATAAACAAATCTCCCCTGAACATCAGGCTATTCTGTCGAAACGTCTTGAGTCTGTCGAA
TCCGATATTCGTCTTGAGCAGAATACCATGGATATGACCCGAATCGATGCGCGCAAGATG
CAGATGACGGGCGATCTGATTATGAAGAACTCGGTCACGGTCGGTGGTATTGCAGGGGCG
TCCGGGCAGTACGCCGCTACTCAGGAACGTTCCGAGCAGCAAATTAGCCAGGTGAATAAC
CGGGTTGCCAGCACCGCATCGGACGAAGCCCGTGAAAGTTCACGTAAATCGACCAGCCTG
ATTCAGGAAATGCTGAAAACAATGGAGAGCATTAACCAGTCGAAAGCATCCGCACTCGCT
GCTATCGCAGGCAATATTCGCGCTTAATCTGAAAGGTCATCTATACGCCATCATGGGTGT
GATTTAATCGCGCTCCTGATGGCGAACTGGGGATATT**ATG**CTTAATATTCAAAATTATTC
CGCTTCTCCTCATCCGGGGATCGTTGCCGAACGGCCGCAGACTCCCTCGGCGAGCGAGCA
CGTCGAGACTGCCGTGGTACCGTCTACCACAGAACATCGCGGTACAGATATCATTTCATT
ATCGCAGGCGGCTACTAAAATCCACCAGGCACAGCAGACGCTGCAGTCAACGCCACCGAT
CTCTGAAGAGAATAATGACGAGCGCACGCTGGCGCGCCAGCAGTTGACCAGCAGCCTGAA
TGCGCTGGCGAAGTCCGGCGTGTCATTATCCGCAGAACAAAATGAGAACCTGCGGAGCGC
GTTTTCTGCGCCGACGTCGGCCTTATTTAGCGCTTCGCCTATGGCGCAGCCGAGAACAAC
CATTTCTGATGCTGAGATTTGGGATATGGTTTCCCAAAATATATCGGCGATAGGTGACAG
CTATCTGGGCGTTTATGAAAACGTTGTCGCAGTCTATACCGATTTTTATCAGGCCTTCAG
TGATATTCTTTCCAAAATGGGAGGCTGGTTATTACCAGGTAAGGACGGTAATACCGTTAA
GCTAGATGTTACCTCACTCAAAAATGATTTAAACAGTTTAGTCAATAAATATAATCAAAT
AAACAGTAATACCGTTTTATTTCCAGCGCAGTCAGGCAGCGGCGTTAAAGTAGCCACTGA
AGCGGAAGCGAGACAGTGGCTCAGTGAATTGAATTTACCGAATAGCTGCCTGAAATCTTA
TGGATCCGGTTATGTCGTCACCGTTGATCTGACGCCATTACAAAAAATGGTTCAGGATAT
TGATGGTTTAGGCGCGCCGGGAAAAGACTCAAAACTCGAAATGGATAACGCCAAATATCA
AGCCTGGCAGTCGGGTTTTAAAGCGCAGGAAGAAAATATGAAAACCACATTACAGACGCT
GACGCAAAAATATAGCAATGCCAATTCATTGTACGACAACCTGGTAAAAGTGCTGAGCAG
TACGATAAGTAGCAGCCTGGAAACCGCCAAAAGCTTCCTGCAAGGATAACAGAAGAGGAT
ATTAATA**ATG**GTTACAAGTGTAAGGACTCAGCCCCCCGTCATAATGCCAGGTATGCAGAC
CGAGATCAAAACGCAGGCCACGAATCTTGCGGCGAATCTTTCCGCAGTCAGAGAAAGTGC
CACAGCGACGCTGTCAGGGGAAATTAAAGGCCCGCAACTGGAAGATTTTCCCGCGCTGAT
CAAACAGGCGAGTCTGGATGC

SEQ ID NO:15

FIG. 28

GAGCTCAGCAACGTGTCGAAAGCCTGTAAAATCATGGGCGTCTCGCGCGATACGTTTTAC
CGTTATCGTGAACTGGCCGATGAAGGCGGCGTTGATGCGCTGATAAATCGTAGTCGCCGC
GTACCTAACCTTAAGAACCGTACCGATGAGGCAACTGAGCAAGCTGTTGTTGATTATGCC
GTTGCGTTCCCGGCCCATGGTCAGCACCGAACTGCGCAAACAGGACGTTTTTATCTCCGG
TAGTGATGTCCATTCCGTCTGGCTGCGCACAACCTTGAGAACTTCAAAAAACGCCTGAAA
GCGCTGGAAGAAAAAGTGGCCCGCGATGGCATTGAACTGACTGCCAGATCGCCGCGCTGG
AGCGTAAAGCCAGTGATGATGAAGCCTGTGGTGAGATTGAAACCGTTCATCCGGGATATC
TGGGGTCACAGGACACGTTCTACGTGGGCAACCTGAAAGGCGTTGGGCGAATCTATCAGC
AGACGTTCGTTGATACATACTCGAAGGTGGCTCACTGCAAGCGCTATATCACCAAAACGC
CGATTACAGCGGCTGATTTGCTGAATGATCGTGTACTGCCGTTTATGAGTCTCAGGGCCT
GCCGATGCTAAGGATACTGACAGACAGGGGTACAGAATATTGCGGCAAAGTGGAACATCA
TGATTATCAGCTTTATCTGGAGATAAATGACATCGAACACACGAAAACGAAGGCGATGTC
CCCGCAGACCAATGGCATCTGCGAGCGGTTCCATAAAACGATACTGAACGAATTTTATCA
GGTGACGTTCCGCAAAAAGTTATATGGCGATTTTGATACATTACAATCGGATCTTGATGA
ATGGCTGGTTCACTATAATAATGAGCGAACCCATCAGGGAAAAATGTGCTGTGGCCGGAC
GCCGATGGAAACGTTACTTGATGGAAAACGCATCTGGTCTGAGAAGAATTTAAGCCAGAT
GTAATCTGACAGATACCTGTATAAATAACCGGTAACTGTCAGATCAGGTCTGAGCTAATA
CAACTAATTGTATGTTATTTGTCGTTTATTGCTAAATATATATCGTTAATTGAAGGCTTG
ATGCGTGTGTCTGCGTTAATCTCTTTTCATTGTGCTGTAAATTAGGCAGTGGAATATGTT
TAATATCCGCAATACACAACCTTCTGTAAGTATGCAGGCTATTGCTGGTGCAGCGGCACC
AGAGGCATCTCCGGAAGAAATTGTATGGGAAAAAATTCAGGTTTTTTTCCCGCAGGAAAAT
TACGAAGAAGCGCAACAGTGTCTCGCTGAACTTTGCCATCCGGCCCGGGGAATGTTGCCT
GATCATATCAGCAGCCAGTTTGCGCGTTTAAAAGCGCTTACCTTCCCCGCGTGGGAGGAG
AATATTCAGTGTAACAGGGATGGTATAAATCAGTTTTGTATTCTGGATGCAGGCAGCAAG
GAGATATTGTCAATCACTCTTGATGATGCCGGGAACTATACCGTGAATTGTCAGGGGTAC
AGTGAAGCACATGACTTCATCATGGACACAGAACCGGGAGAGGAATGCACAGAATTCGCG
GAGGGGGCATCCGGGACATCCCTCCGCCCTGCCACAACGGTTTCACAGAAGGCAGCAGAG
TATGATGCTGTCTGGTCAAATGGGAAAGGGATGCACCAGCAGGAGAGTCACCCGGCCGCG
CAGCAGTGGTACAGGAA**ATGCGTGATTGCCTGAATAACGGCAATCCAGTGCTTAACGTGG**
GAGCGTCAGGTCTTACCACCTTACCAGACCGTTTACCACCGCATATTACAACACTGGTTA
TTCCTGATAATAATCTGACCAGCCTGCCGGAGTTGCCGGAAGGACTACGGGAGCTGGAGG
TCTCTGGTAACCTACAACTGACCAGCCTGCCATCGCTGCCGCAGGGACTACAGAAGCTGT
GGGCCTATAATAATTGGCTGGCCAGCCTGCCGACGTTGCCGCCAGGACTAGGGGATCTGG
CGGTCTCTAATAACCAGCTGACCAGCCTGCCGGAGATGCCGCCAGCACTACGGGAGCTGA
GGGTCTCTGGTAACAACCTGACCAGCTGCGCGCGCTGCCGTCAGGACTACAGAAGCTGTG
GGCCTATAATAATCGGCTGACCAGCCTGCCGGAGATGTCGCCAGGACTACAGGAGCTGGA
TGTCTCTCATAACCAGCTGACCCGCCTGCCGCAAAGCCTCACGGGTCTGTCTTCAGCGGC
ACGCGTATATCTGGACGGGAATCCACTGTCTGTACGCACTCGTGACAGGCTCTGCGGACA
TCATTGGCCATTCAGGCATCAGGATACACTTCGATATGGCGGGGCCTTCCGTCCCCGGGA
AGCCCGGGCACTGCACCTGGCGGTCGCTGACTGGCTGACGTCTGCACGGGAGGGGGAAGC
GGCCCAGGCAGACAGATGGCAGGCGTTCGGACTGGAAGATAACGCCGCCGCCTTCAGCCT
GGTCCTGGACAGACTGCGTGAGACGGAAAACTTCAAAAAAGACGCGGGCTTTAAGGCACA
GATATCATCCTGGCTGACACAACTGGCTGAAGATGCTGCGCTGAGAGCAAAAACCT
TTGCCATGGCAACAGAGGCAACATCAACCTGCGAGGACCGGGTCACACATGCCCTGCACC
AGATGAATAACGTACAACTGGTACATAATGCAGAAAAAGGGGAATACGACAACAATCTCC
AGGGGCTGGTTTCCACGGGGCGTGAGATGTTCCGCCTGGCAACACTGGAACAGATTGCCC
GGGAAAAAGCCGGAACACTGGCTTTAGTCGATGACGTTGAGGTCTATCTGGCGTTCCAGA
ATAAGCTGAAGGAATCACTTGAGCTGACCAGCGTGACGTCAGAAATGCGTTTCTTTGACG
TTTCCGGCGTGACGGTTTCA
GACCTTCAGGCTGCGGACGTTCAGGTGAAAACCGCTGAAAACAGCGGGTTCAGTAAATGG
ATACTGCAGTGGGGGCCGTTACACAGCGTGCTGGAACGCAAAGTGCCGGAACGCTTTAAC
GCGCTTCGTGAAAAGCAAATATCGGATTATGAAGACACGTACCGGAAGCTGTATGACGAA
GTGCTGAAATCGTCCGGGCTGGTCGACGATACCGATGCAGAACGTACTATCGGAGTAAGT
GCGATGGATAGTGCGAAAAAAGAATTTCTGGATGGCCTGCGCGCTCTTGTGGATGAGGTG
CTGGGTAGCTATCTGACAGCCCGGTGGCGTCTTAACTGAGCACGATATTCTCCGCACCAG
GCGAATGTGGTGCGGTGAACAAAGATATTCCTTGGACAAACAACATGAGACAGCACTGAT
GATGCACAGGTGAAACAGGGGAGACTTCTTCAGTCAGGGCGTACGCAACTCAACCTTTTC
GACGATACGCGCC/

SEQ ID NO:16

FIG. 29

PTO/PCT Rec'd 19 OCT 1998

PATENT
ATTORNEY DOCKET NO: 00786/292002

COMBINED DECLARATION AND POWER OF ATTORNEY

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name,

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled SALMONELLA SECRETED PROTEINS AND USES THEREOF, the specification of which

☐ is attached hereto.
■ was filed on May 14, 1998 as Application Serial No. 09/068,804 and was amended on ____________.
■ was described and claimed in PCT International Application No. PCT/US96/18504 filed on 14 November 1996 and was amended on ________.

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose all information I know to be material to patentability in accordance with Title 37, Code of Federal Regulations, §1.56.

I hereby claim foreign priority benefits under Title 35, United States Code, §119 of any foreign application(s) for patent or inventor's certificate or of any PCT international application(s) designating at least one country other than the United States of America listed below and have also identified below any foreign application for patent or inventor's certificate or any PCT international application(s) designating at least one country other than the United States of America filed by me on the same subject matter having a filing date before that of the application(s) of which priority is claimed:

COUNTRY	APPLICATION NO.	FILING DATE	PRIORITY CLAIMED
			☐ Yes ☐ No

I hereby claim the benefit under Title 35, United States Code, §120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of Title 35, United States Code, §112, I acknowledge the duty to disclose all information I know to be material to patentability as defined in Title 37, Code of Federal Regulations, §1.56(a) which became available between the filing date of the prior application and the national or PCT international filing date of this application:

U.S. SERIAL NO.	FILING DATE	STATUS
60/006,733	14 November 1995	■ Pending ☐ Issued ☐ Abandoned
PCT/US96/18504	14 November 1996	■ Pending ☐ Issued ☐ Abandoned

I hereby appoint the following attorneys and/or agents to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith: Janis K. Fraser, Reg. No. 34,819; Ingrid A. Beattie, Reg. No. P-42,306; John W. Freeman, Reg. No. 29,066; J. Peter Fasse, Reg. No. 32,983; Y. Rocky Tsao, Reg. No. 34,053; Eldora L. Ellison, Reg. No. 39,967; Timothy A. French, Reg. No. 30,175; and Anita L. Meiklejohn, Reg. No. 35,283.

8

Address all telephone calls to Janis K. Fraser, Ph.D. at telephone number 617/542-5070.

Address all correspondence to Janis K. Fraser, Ph.D., Fish & Richardson P.C., 225 Franklin Street, Boston, MA 02110-2804

COMBINED DECLARATION AND POWER OF ATTORNEY CONTINUED

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patents issued thereon.

1-00 Full Name of Inventor: Samuel I. Miller

Inventor's Signature: Samuel I. Miller Date: Oct 13, 1998

Residence Address: ~~5322 N.E. 86th Street, Seattle, Washington 98115~~ 6033 45th Ave NE

Citizen of: United States of America (SEATTLE) Seattle WA

Post Office Address: ~~5322 N.E. 86th Street, Seattle, Washington 98115~~

6033 45th Ave NE
Seattle, WA

306971.B11